View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by ScholarBank@NUS

CONSTITUTIVE EMERGENCE: A NON-REDUCTIVE SOLUTION TO THE MIND-BODY PROBLEM

WONG SOO LAM

(BA PHILOSOPHY – UNIVERSITY OF LONDON)

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF PHILOSOPHY

NATIONAL UNIVERSITY OF SINGAPORE

2007

ACKNOWLEDGEMENT

My utmost gratitude goes to my thesis supervisor Assoc. Prof. Cecilia Lim. I benefitted tremendously from her excellent insights into the subject matter and her unfailingly enthusiastic guidance.

My heartfelt thanks go to administrator Ms Devi Asokan for consistently offering outstanding support and providing helpful information.

I am extremely grateful to the Department of Philosophy of NUS in giving me this opportunity to realise my dream in pursuing higher studies in my chosen field.

Last but not least, I offer my deepest appreciation to my understanding and supportive wife Catherine Neo, who has to put up with my frequent late nights and occasional philosophical ramblings.

CONTENT

Chapter Three: The Concept of Emergence and the Mind-Body Problem

Chapter Four: Constitutive Emergence: A Non-Reductive Approach To The Mind-Body Problem

SUMMARY

The mind-body problem has two essential aspects – the problem of mental causation and the problem of consciousness. Substance or property dualism is able to address the latter but not the former aspect adequately. Reductive physicalism is able to address the former but not the latter aspect adequately. Non-reductive physicalism, which aims to address both aspects of the mind-body problem adequately, has established itself as the orthodox position. However, Jaegwon Kim asserts that non-reductive physicalism is an inherently unstable position in holding both the mutually incompatible non-reductivism thesis and physicalism thesis. This argument introduces a dilemma in non-reductive physicalism, which leads to its inevitable collapse into dualism, epiphenomenalism or reductive physicalism. To avoid this collapse, I propose a version of non-reductive physicalism based on the contentious notions of constitution and emergence. In my view, constitution is neither the identity of parts and wholes nor the co-location of two distinct objects. And emergence is neither a strong ontological thesis which affirms downward causation nor a weak epistemological thesis which denies relational properties. Rather, constitution and emergence are non-reductive, part-whole notions with emphasis on the configuration of parts and relational properties of wholes. I argue that this view, which I call constitutive emergence, is able to hold on to both the non-reductivism and physicalism theses without collapsing into dualism, epiphenomenalism or reductive physicalism.

CHAPTER ONE

NON-REDUCTIVE PHYSICALISM AND THE MIND-BODY PROBLEM

Introduction

The Mind-Body Problem (MBP) has two essential aspects. Substance or property dualism, in holding that mind and body are separate and distinct substances or properties, faces a problem of explaining how the non-physical substance or properties of the mind interact with the physical substance or property of the body and vice versa.[1] In other words, dualism cannot adequately explain how the non-physical mental states can have causal power over physical states and vice versa. This first aspect of the MBP is known as the problem of mental causation. Reductive physicalism, in holding that mind is nothing over and above the brain, or that mental states are nothing over and above brain states, faces a problem of explaining how our inherently private mental processes are nothing over and above our publicly observable brain processes. In other words, reductive physicalism cannot adequately explain the qualitative differences between the mind and the brain, or between mental states and brain states. This second aspect of MBP is known as the problem of consciousness.[2] Put simply, dualism provides a solution to the problem of consciousness but faces the problem of mental causation;

[1]In this paper, substance dualism refers only to Descartes' interactionism between minds and bodies, which faces the problem of mental causation. It does not refer to Spinoza's and Leibniz's parallelism, where there are no interactions between minds and bodies. Note that though parallelism does not face the problem of mental causation, it does collapse into some form of epiphenomenalism discussed later.

[2] Throughout this paper, I take consciousness to mean broadly the subjective nature of our experiences or our first-person perspective. The problem of consciousness, then, refers to the inadequacy of reductionist theories of mind in accounting for the subjective nature of our experiences or our first-person perspective.

reductive physicalism avoids the problem of mental causation but faces the problem of consciousness.

In this study, I seek a solution to MBP that attempts to solve both the problem of consciousness and the problem of mental causation simultaneously. And I have to state up front that my task here is to derive a conceptual solution to a conceptual problem, and not an empirical solution to an empirical question. For my research objective, I do not attempt to answer the question 'How does the biological brain give rise to conscious states?' nor do I aim to survey the latest cognitive or neurological findings to shed light on this question. Instead, I would like to provide an answer to the question 'How to conceive the mind in relation to the body?' by drawing from the relevant insights of various philosophers to that effect. While focused on answering the conceptual question 'How to conceive the mind in relation to the body?' I do not deny that the empirical question 'How does the biological brain give rise to conscious states?' forms an important part of MBP. But the former question can be answered by building on a string of related concepts formulated by some philosophers while the latter requires further empirical studies. Moreover, I believe that I can work towards an adequate answer to the former question without touching much on the latter.

An adequate solution to the MBP requires us to solve both the problems of mental causation and consciousness simultaneously. Two principles serve to guide us in deriving an adequate solution. The first principle, Ockham's razor, requires us not to assume plurality without necessity. In the context of MBP, it can mean that we need not

postulate non-physical substances or properties unnecessarily to explain the phenomenon of conscious experiences. The second principle, saving appearances, requires us to account for how things seem like to us. In the context of MBP, it can mean that we need to account for the subjective nature of our mental states and not to simply explain it away. In the light of these two principles, it would seem that an adequate solution to MBP lies in the middle road between dualism and reductive physicalism, that is, towards some form of non-reductive physicalism (NRP). To understand what NRP entails, we must first have a working definition of physicalism, reductionism before we can examine NRP itself.

Physicalism

Jaegwon Kim, in his article '*The Mind-Body Problem at Century's Turn'*, provides us with a useful working definition of physicalism as follows:

"The core of contemporary physicalism is the thesis that all things that exist in the world are bits of matter and structures aggregated out of bits of matter, all behaving in accordance with the laws of physics. This metaphysical thesis has a companion epistemological thesis, the claim that all phenomena of the world can be physically explained if they can be explained at all." (Kim 2004, P.129)

If the physicalism thesis is true, then the problem of mental causation arises. For how do minds fit into a world where only matter and aggregates of matter are all that exist? In

his *'Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalist'*, Kim makes explicit the causal implications of the physicalism thesis, which can be expressed in two principles. The first he calls the principle of the Causal Closure of the Physical Domain:

Principle Causal Closure of the Physical Domain: If a physical event has a cause at t, then it has a physical cause at t.

Note that this principle only states that all physical events have physical causes. It does not rule out non-physical events that may have non-physical causes. To rule out non-physical causes, we require the second principle of Causal Exclusion:

Principle of Causal Exclusion: If an event e has a sufficient cause c at t, then no event at t distinct from c can be the cause of e (unless this is a genuine case of causal over-determination).

Combined with the first principle, the second principle implies the following Principle of the Causal Exclusion of the Physical Domain:

Principle of the Causal Exclusion of the Physical Domain: If a physical event e has a sufficient physical cause c at t, then no event at t distinct from c can be the cause of event e.

This principle rules out non-physical causes of physical events. And if non-physical causes are ruled out, then reductionism of mental causes follows.

Reductionism

In the context of MBP, reductionism is the idea that mental events are reducible to (metaphysical reduction) or reductively explainable in terms of (epistemological reduction) physical events. Three forms of reductionism are consistent with the principles of causal closure and causal exclusion of the physical domain. The first is known as the mind-brain identity theory, where the mind is (nothing over and above) the brain or where mental states are (nothing over and above) brain states. Here, minds and mental states do not have causal powers on their own and their causal powers are identical to those of the brain and brain states. The second and more extreme form of reductionism is known as eliminative physicalism, which denies that minds and mental states exist. Here, only brains and brain states exist; minds and mental states are only fictitious constructions. The third is known as epiphenomenal dualism that affirms the existence of minds and mental states but denies that they have any causal powers. Here, causally impotent mental states simply tag along with the causally potent brain states. Some philosophers would find all these forms of reductionism inadequate in addressing the problem of consciousness and argues for mental realism, the view that minds and mental states are real and causally relevant. As the problem of mental causation has rendered dualism an implausible position, most of these thinkers would want to keep both the thesis of physicalism and mental realism, coming up with various forms of NRP.

Non-Reductive Physicalism

Kim discussed three forms of NRPs and shows how all of them eventually collapse into physical reductionism. The first form is Donald Davidson's Anomalous Monism, which can be summarised as follows:

1. Mental events enter into causal relations with physical events.
2. Causal relations must be backed by laws (or must exhibit lawful regularities).
3. There are no psychological or psychophysical laws.

---

4. Hence, any causal relation involving a mental event must instantiate a physical law; or any causal relation involving a mental event has a physical description and falls under a physical event.
5. And hence, a mental event is also a physical event; or a mental property is also a physical property

According to Anomalous Monism, events are causes or effects only if they instantiate physical laws or an event's mental properties can only figure in causal relations only when they are also physical properties. This implies that only physical events or properties can figure as causes or effects, and that mental events or properties are causally or explanatorily impotent. Hence, Anomalous Monism has collapsed into a form of reductionism – namely epiphenomenal dualism.

The second form is the Supervenience Argument. Donald Davidson, in his article *'Mental Events'*, describes the notion of supervenience as follows:

"Although the position I describe denies that there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respects, or that an object cannot alter in some mental respect without altering in some physical respect. Dependence or supervenience of this kind does not entail reducibility through law or definition: if it did, we could reduce moral properties to descriptive, and this there is good reason to believe cannot be done." (Davidson 1980, P.214)

Kim observed that this notion of supervenience has two requirements. First, the relation between mental (supervenient property) and physical (subvenient property) characteristics is non-reductive. Second, the relation is one of dependence. That is, mental (supervenient property) is dependent on or determined by physical (subvenient property) characteristics. And Kim identified a problem with the supervenience relation. If the relation is weak enough to be non-reductive, it tends to be too weak to serve as a dependence relation. For it is compatible with co-relating but independent mental and physical characteristics held by dualism. If the relation is strong enough to support dependence, then it is strong enough to imply reductionism. For if mental characteristics are completely dependent on or determined by physical characteristics, then the mental characteristics have no causal powers on their own. The dilemma faced by the

supervenience relation is this: either it is too weak and collapses into dualism or it is too strong and collapses into reductive physicalism.

The third form is the Multiple Realisation Argument, which can be formulated as follows: for any mental state, there is an indefinite number of nomologically possible physical states such that, though each of physical states realises or implements the mental state, none of them are co-extensive with it. Hence, a mental state cannot be reducible to the disjunctive physical realisers. Kim concedes that although a global or general reduction of mental states to their physical realisers (also known as the type identity theory) is impossible, a local or specific reduction is still possible. For example, we cannot reduce a mental state m1 to p1 for all cases but we can reduce mental state m1 to p1 in organism 1, p2 in organism 2, and p3 in organism etc. And so m1 can be co-extensive with p1 in organism 1, and p2 in organism 2 etc. In this way, the causal power of m1 can be reducible to p1 in organism 1 etc. Hence, the Multiple Realisation Argument has collapsed into a form of reductionism – namely the token identity theory.

According to Kim, NRP generally affirms the following theses:

1. Mental properties are not reducible to physical properties. (Irreducibility)
2. Mental properties are causal properties. (Mental Realism)
3. Mental events cause physical events. (Downward Causation)
4. If a physical event has a cause at t, then it has a physical cause at t. (Causal Closure of the Physical Domain)

Thesis (1) rejects all forms of mind-brain identity theory. Theses (2) and (3) reject epiphenomenal dualism. Note that thesis (4) is compatible with theses (1) to (3). More importantly, although it states that physical events must have physical causes, it does not deny that physical events can have mental causes as well. Hence, thesis (4) leaves room for theses (1) to (3).

Kim argues that NRP is unstable. Suppose that a certain event c causes a physical event e in virtue of its mental property. Thesis (4) requires that a physical event must also have a physical cause. So we may assume that event c causes event e in virtue of its physical property as well. The following question now arises: what is the relationship between these two causes, one mental and the other physical? Kim discussed two possibilities: first, we can say that each is only a partial and necessary cause, and the two together make the complete and sufficient cause. This violates thesis (4) by saying that part of event c goes outside the physical domain. Second, we can say that each is an independent and sufficient cause, and this means event e is over-determined. This violates thesis (4) too by saying that if the physical cause had not occurred, the mental cause would have caused the effect by itself. Hence, a more plausible answer is required for the question: how are the mental cause and the physical cause of the single effect related to each other?

One plausible option is to identify the mental cause with the physical cause, so that there is only one cause and not two. The identification of mental properties with physical

properties is the central tenet of the mind-brain identity theory. Here, theses (1), (2) and (3) are rejected. And if NRP accepts this, it collapses into a form of reductive physicalism. Another plausible option is to say that there is only one physical cause that can be described in two different ways – mentally and physically. This position is known as conceptual dualism which accepts ontological reduction but rejects epistemological reduction of the mental to the physical. Here, theses (2) and (3) are rejected. And if NRP accepts this, it collapses into a form of epiphenomenal dualism. Hence, as Kim noted, NRP is an unstable position. Are there any other answers NRP can offer to the question: what is the relationship between the mental cause and the physical cause of a single effect? Or put another way: how are the mental cause and the physical cause of a single effect related to each other?

Constitutive Emergence as Non-Reductive Physicalism

In what follows, I would like to explore the notions of constitution and emergence as a form of NRP. Constitution is best conceived of as a mereological or part-whole relation. Some philosophers believe that constitution is an identity relation. This is the view that the whole is not greater than the sum of its parts, or that the whole is identical with the sum of its parts. In the context of MBP, it implies that if the brain constitutes the mind, then the brain is identical with the mind. This view is not able to account for qualitative differences between the whole and the sum of its parts. Or in the context of MBP, it is not able to account for the qualitative differences between the brain and the mind. Other philosophers believe that constitution is not an identity relation and that the whole is

more than the sum of its parts. In the context of MBP, it implies that if the brain constitutes the mind, then the brain is not identical with the mind. This view is not able to account for there being two distinct things in the same place at the same time. Or in the context of MBP, it is not able to account for there being separate and distinct brain and mind, existing in the same place at the same time. In chapter two, I would discuss the notion of constitution in greater detail, argue for the view that that constitution is a mereological relation without identity, and examine its implications for MBP.

Emergence is the idea that the whole is more than the sum of its parts. When objects reach a certain level of organisational complexity, new properties emerges at the level of wholes that are not possessed by its parts. These new properties possessed by the whole have causal powers over the properties possessed by its parts. In the context of MBP, it implies that the mind emerges from the organisational complexity of the brain and that the emergent mental properties have causal powers over basal physical properties. This is the stronger view of emergence that postulates configurational forces and affirms downward causation. More generally, it accepts theses (1) to (3). There is a weaker view of emergence that does not postulate configurational forces and denies downward causation. More generally, it accepts theses (1) and (2) but not (3). I agree with the latter view that it is possible to affirm emergence without accepting configurational forces and downward causation. In chapter three, I would discuss the notion of emergence in greater detail, argue for the view that emergent properties are simply the ways objects are, and examine its implications for MBP. It is important to note that the notions of constitution and emergence, when they are construed in certain ways, are compatible with one

another. In fact, they are actually the converse of one another. While constitution takes the bottom-up (parts to whole) approach, emergence takes the top-down (whole to parts) approach.

In this study, I aim to show that NRP in the form of constitutive emergence presents one of the most promising options compared to those discussed above in resolving the MBP. It wields the Ockham's Razor – the first guiding principle – by not postulating non-physical substances and properties to solve the problem of consciousness. And it attempts to save appearances – the second guiding principle – by not reducing mental states to merely brain states to solve the problem of mental causation. As noted by Kim, the central challenge faced by NRP is its inherent instability. In affirming both non-reductivism (theses 1 to 3) and physicalism (thesis 4), NRP risks collapsing into either dualism or reductive physicalism. It is my aim to show that NRP, construed as constitutive emergence, is resilient against either collapse. In the concluding chapter, I would summarise the constitutive emergence approach to the MBP by integrating both the notions of constitution and emergence, and clarifying its ontological and epistemological commitments.

Reference:

Davidson, Donald (1970) 'Mental Events' in *Experience and Theory* ed. L. Foster & J. Swanson (University of Massachusetts Press)

Davidson, Donald (1993) 'Thinking Causes' in *Mental Causation* ed. John Heil and Alfred Mele (Oxford: Clarendon Press) P.3-19

Kim, Jaegwon (1993) 'The Non-Reductivist's Troubles with Mental Causation', in *Mental Causation* ed. John Heil and Alfred Mele (Oxford: Clarendon Press) P.189-211

Kim, Jaegwon (1994) 'The Myth of Non-reductive Materialism', in *The Mind-Body Problem: A Guide to the Current Debate* ed. Richard Warner and Tadeusz Szubka (Oxford University Press) P.242-261

Kim, Jaegwon (2001) 'The Two Mind-Body Problems for the Physicalist', in *Physicalism and its Discontents* ed. Carl Gillet and Barry Loewer (Cambridge University Press) P.271-284

Kim, Jaegwon (2004) 'The Mind-Body Problem at Century's Turn', in *The Future for Philosophy* ed. Brian Leiter (Oxford: Clarendon Press) P.129-153

CHAPTER TWO

THE CONCEPT OF CONSTITUTION AND THE MIND-BODY PROBLEM

Introduction

Constitution is a key conceptual component of non-reductive physicalism. It is the idea that the whole is made up of its parts. I would first examine the problem of material constitution and some of the solutions proposed by various philosophers. The solutions are divided into mainly two different camps – one advocating the identity between the whole and its constituting parts and another advocating the distinction between the whole and its constitution parts. This is followed by a discussion of the various versions of identity thesis and their difficulties. Next, I would present Lynn Rudder Baker's thesis of constitution, which I have some sympathy for and highlight a few difficulties. Then, I will argue for a modified constitution thesis and conclude this chapter by examining how this modified constitution thesis can help shed light on the mind-body problem.

The Problem of Material Constitution

The problem of material constitution arises from analysing relationship between parts and wholes. In particular, it is the problematic relationship between a material object and the material parts that constitute it. The 'Constitution is Identity' view holds that the material object is identical to the material parts that constitute it. According to this view, it is not

possible for two objects to exist in the same place at the same time and hence the material object and its constituting parts must be identical. The 'Constitution is Not Identity' view holds that the material object is distinct from the material parts that constitute it. According to this view, the material object and its constituting parts cannot be identical because there could be at least one property they do not share and this implies that it is possible for two objects (the material object and the material parts) to exist in the same place at the same time. The defenders of the 'Constitution is Identity' view need to show that the differences between the material object and its constituting parts are merely apparent. And the notions of strict identity, contingent identity, relative identity, Abelardian predicates will be discussed to evaluate their success in doing so. The defenders of the 'Constitution is Not Identity' need to show that their view does not necessarily imply two entities existing in the same place at the same time although there are some who accepts this consequence. And the notion of accidental sameness and Lynn Rudder Baker's notion of constitution will be used to represent this view and its success is assessed accordingly.

Constitution is Identity Theses

Constitution as Strict Identity

In his paper *'Composition as Identity'*, Peter van Inwagen describes mereology as a theory about composition – the relationship of parts and wholes, and goes on to discuss the view of David Lewis presented in *'Parts of Classes'*. Lewis' central claim is that

Mereology is ontologically innocent. That is, in accepting fusions, mereology requires us to accept only what is identical with what we have already accepted. And Mereology asks us to accept nothing more than the fusions of what we already accept. This point can be expressed as either 'y is a fusion of xs' or 'y is composed of xs'. When we accept y, we are only accepting the totality of xs that compose y and nothing else. In other words, 'Composition is a kind of identity'. The following quote is a summary of Lewis' (unqualified) view:

"I say that composition – the relation of part to whole, or, better, the many-one relation of the many parts and their fusion – is like identity. The 'are' of composition is, so to speak, the plural form of the 'is' of identity. Call this the thesis of 'Composition as Identity'. It is virtue of this thesis that mereology is ontologically innocent; it commits us only to things that are identical, so to speak, to what we were committed to before." (Lewis 1991, P.82)

Lewis's view can be made clearer by examining various other quotes from his 'Parts of Classes':

1. *"The fusion of the xs just is the xs."* (Lewis 1991, P.81)
2. *"The xs are the fusion of the xs."* (Lewis 1991, P.81)

I read statements (1) and (2) as saying that the whole is nothing more than its aggregate parts. The 'is' and 'are' refer to identity relations.

3. *"Commit yourself to the existence of the xs all together or one at a time, it's the same commitment either way."* (Lewis 1991, P.81)
4. *"The whole is the many parts counted as one thing."* (Lewis 1991, P.83)

I read statements (3) and (4) as saying that the whole and its aggregate parts are to be counted as a single entity. They are one and the same thing.

5. *"Take the xs together or take them separately, the xs are the same portion of Reality either way."* (Lewis 1991, P.81)
6. *"If you draw up an inventory of Reality…it would be double counting to list the fusion of the xs and also list the xs."* (Lewis 1991, P.81)

I read statements (5) and (6) as saying that the aggregate of parts are all that exists and the whole does not has a separate and distinct existence from the aggregate of its parts.

From these statements, Lewis argues that mereology is ontologically innocent because the whole is 'nothing more than', 'just the summation of', 'not separate and distinct from' its aggregate parts. The whole is equivalent without residue to the composition of its parts. Again in short, 'composition is identity'.

Van Inwagen observed that Lewis eventually qualified his identity claim.

"...even though the many and the one are the same portion of Reality, and the character of the portion is given once and for all whether we take it as many or take it as one, still we do not really have a generalised principle of indiscernibility of identicals. It does matter how you slice it – not to the character of what's described, of course, but to the form of the description. What's true of the many is not exactly true of the one. After all, there are many while it is one." (Lewis 1991, P.87)

In this passage, Lewis seems to admit that strict identity (as expressed by the principle of indiscernibility of identicals) between the whole and its aggregate parts may not be true, as 'what's true of the many is not exactly true of the one'. Nonetheless, in saying that 'there are many while it is one', he seems to insist that the whole and its aggregate parts are still identical in some sense. Lewis elaborates:

"Mereological relations…are strikingly analogous to ordinary identity. So striking is this analogy that it is appropriate to mark it by speaking of mereological relations – the many -one relation of composition, the one-one relations of part to whole and overlap – as kinds of identity. Ordinary identity is the special limiting case of identity in the broadened sense." (Lewis 1991, P.83)

As remarked by Van Inwagen, this passage seems to say that our vocabulary does not reflect the striking analogy between composition and identity and it is hence appropriate to expand the meaning of 'identity' to cover both composition and identity. In 'Parts of Classes' (P.85), Lewis tells us that the analogy has many aspects:

a. ‘Just as it is redundant to say that x and y exist when x is identical to y, so it is redundant to say that x and the ys exist, when x is a fusion of the ys.’
b. ‘Just as given that x exists, it is automatically true that something identical with x exists, so given that the xs exists, it is automatically true that a fusion of the xs exists.’
c. ‘Just as there cannot be two things both of which are identical with x, so there cannot be two things both of which are fusions of xs.’
d. ‘Just as to fully describe the x is to fully describe the object that is identical with x, so to fully describe xs is fully to describe their fusion.’
e. ‘Just as x and y must occupy the same region of space-time if the former is identical with the latter, so x and the ys must occupy the same region of space-time if the former is the fusion of the latter.

The following quote Van Inwagen’s summary of Lewis’ (qualified) view:

“Since composition and identity are analogous in these five respects, it is philosophically appropriate to expand the meaning of ‘identity’ to cover both composition and identity, and to regard what used to be called ‘identity’ as the ‘special limiting case’ of what is now called ‘identity.” (Van Inwagen 1994, P.217)

Van Inwagen commented that the analogy between composition and identity is so weak as really not to be much of an analogy at all. One may still insist that if the whole and its

aggregate parts are indeed identical, then they should not be discernible. If mereology is a correct theory of composition, Lewis's composition as identity thesis would appear to be correct – and otherwise not. And if the 'composition as identity thesis' is correct, then mereology is ontologically innocent. As I will make clear later on, one reason why the problem of material constitution arises is because of the failure to distinguish between material parts and ways the material parts are. If x and the fusions of xs are identical, I assume that they are strictly the same object which must have all properties in common, as required by Leibniz's criterion of identity. However, as we will see below, it is possible for xs and the fusion of xs to be qualitatively different. If this is true, then mereology may not be ontologically innocent as Lewis claims it to be.

Constitution as Contingent Identity

Allan Gibbard, in his article '*Contingent Identity*', developed the popular Lumpl/Goliath story to discuss the problem of material constitution. A sculptor sets out to construct the statue of Goliath in the following way: he sculpts Goliath's upper body from one piece of clay and lower body from another, and then joins the two pieces of clay together to form one new piece of clay which is Goliath. This new piece of play is allowed to harden and he smashes it the next day (perhaps because the sculpture is not good enough), thereby destroying both the new piece of clay and Goliath.

The first question here is whether the piece of clay (Lumpl) is identical with the statue Goliath. Gibbard says yes. Lumpl and Goliath come into existence at the same time

when the two difference pieces of clay (upper and lower body of the statue) come together, they are constituted by the same clay parts, and they occupy the same space at the same time before their destruction. The second question here whether the piece of clay (Lumpl) is necessarily identical with the statue Goliath. Gibbard says no because it is possible that Lumpl continue to exist while Goliath does not, or that Goliath continue to exist while Lumpl does not. In the former case, one can imagine the sculptor moulding Lumpl (the same piece of clay) to another statue Hercules. In the latter case, one can imagine the sculptor re-sculpting Goliath with a piece of bronze instead. Hence, Gibbard says that Lumpl and Goliath are only contingently but not necessarily identical. In her article 'Why Constitution is not Identity?', Lynn Rudder Baker advanced two criticisms against Gibbard's view The first is metaphysical and the second epistemological.

Metaphysically, the constitution as contingent identity view affords no unified account of the relation between persons and bodies or between Lumpl and Goliath. This view holds that x and y are contingently identical only if x and y share all their categorical properties, where categorical properties are those that do not depend on how things are in other worlds. For example, if x and y are contingently identical, then both x and y start to exist at the same time and cease to exist at the same time. Gibbard suggests that a person cease to exist when the body dies. If a person's body continues to exist (perhaps buried) after death, then the person is not contingently identical with the body. And if the body is destroyed (perhaps cremated) after death, then the person is contingently identical with the body. On this view then, whether or nor a person is contingently identical with the body depends on the fate of the body at death, affirmative when the body continues to

exist at death and negative when the body is destroyed at death. Hence, the constitution as contingent identity view affords no unified account of the relation between person and bodies, between Lumpl and Goliath, or between any other things related by constitution. According to Baker, the constitution without identity view can offer a unified account of such relations.

Epistemologically, we would not be justified to know whether x and y are contingently identical while they exist if the constitution as contingent identity view is correct. This is because in order identity of x and y to be contingent, it must be possible that they have different properties; but in order for x and y to be identical at all, they must actually have all their properties in common, including ceasing to exist at the same. Before the demise of x and y however, we are not justified to know whether they will differ or be identical in the future. According to Baker, this problem will not arise on the constitution without identity view, where if x constitutes y now, we can be justified in asserting now that x constitutes y. On the constitution as contingent identity view, where if x is contingently identical to y now, we cannot be justified in asserting now that x is contingently identical to y. Baker's constitution without identity view will be discussed in greater detail later.

Constitution as Relative Identity

In his article *'Constitution is Not Identity'*, Mark Johnston discussed two ways of stating 'constitution is identity'. The two ways are:

a. If y is a paradigm F and x is intrinsically exactly like y then x is an F.

Applying this statement to the statue example: if Goliath is a paradigm statue and Lumpl is intrinsically exactly like Goliath, then Lumpl is a statue.

b. If y is a paradigm F and x is an entity that differs from y in any respect relevant to being an F only minutely, then x is an F.

Applying this statement to the statue example: if Goliath is a paradigm statue and Lumpl is an entity that differs from Goliath in any respect relevant to being a paradigm statue only minutely, then Lumpl is a statue. Johnston argues that (a) could be true but (b) is false, and that the acceptance of (a) requires the acceptance of (b). And since (b) is false, (a) could not be true. I would not want to elaborate on Johnston's argument but instead comment on the success and relevance of these two statements in resolving the problem of material constitution.

In her article 'Why Constitution is not Identity' Lynn Rudder Baker says that it is false to say that if y is an F in virtue of its relational properties, and x is intrinsically like y, then x is an F. Baker's argument against relative identity can be formulated as follows:

Let F = having a specific artistic value.

Let y = Goliath

Let x = Lumpl (a piece of clay)

a. To be relatively identical (or intrinsically exactly alike), both x and y must be F.

b. Goliath (y) has the relational property of 'having a certain artistic value' (F). That is, y is F.

c. Lumpl (x) does not have the relational property of 'having a certain artistic value' (F). That is, x is not F.

---

d. Therefore, Lumpl (x) is not relatively identical to (or not intrinsically exactly like) Goliath (y). That is, $x \neq y$.

The argument for relative identity aims to establish $x = y$ by showing that if $y = F$ and $x = y$, then $x = F$. It failed to establish $x = y$ because $y = F$ but $x \neq F$. If Baker is correct, then relative identity does not succeed in solving the problem of material constitution.

A lot seems to depend on what 'intrinsically exactly like' means in statement (1). If it means 'sharing exactly the same parts as' or 'occupying the same space at the same time as', then perhaps one can say that 'x is identical to y' or simply that 'x is y' instead of the weaker 'x and y are Fs'. This weaker claim of 'y is F and x is F because x is intrinsically exactly like (or only minutely different from) y', where entities x and y are defined in relation to a sortal term F. On this claim, there is a possibility for x to be 'intrinsically exactly like' and yet 'non-intrinsically unlike' y, depending on how the sortal term F is defined. At most, relative identity can only tell us that entities x and y can be classified as an F sort. It tells us whether x and y belong to the same class or kind of things, but stops short of telling us directly whether x is in fact identical to y. My focus here shall be

on whether the xs and the fusion of xs are the same object (x = y, assuming y = fusion of xs), and not whether x and y belong to the same sort, class or kind (if y = F and x = y, then x = F). Hence, I shall put aside relative identity as a relevant solution to the problem of material constitution as I see it.

Abelardian Predicates

In his 'Constitution is Identity', Harold Noonan argues contra Johnston that the thesis of constitution is identity is correct. He uses the notion of Abelardian Predicates to show that the thesis' conflict with Leibniz's criterion for identity is only apparent. According to Noonan, an Abelardian predicate is a predicate whose reference can be affected by the subject term to which it is attached. Accepting Abelardian predicates makes it possible to maintain that Lumpl is identical to Goliath even though it is true that Lumpl has been squeezed into a ball and not destroyed, but false that Goliath might have been squeezed into a ball and not destroyed. This is because one can say that the property denoted by the predicate 'might have been squeezed into a ball and not destroyed' in the true sentence 'Lumpl might have been squeezed into a ball and not destroyed' is not the same as the property as that denoted by that predicate in the false sentence 'Goliath might have been squeezed into a ball and not destroyed.

One way to make sense of Abelardian predicates is through the use of sortal terms. Lumpl and Goliath refer to the same statue (sortal term). The predicate 'might have been squeezed into a ball and not destroyed' is true of the statue qua Lumpl, but it is false of

the statue qua Goliath. This solution to the problem of material constitution suggests that the problem here is linguistic rather than ontological, for there are actually two (or more) ways to describe a material object. Lumpl and Goliath are just two ways of describing the same statue. From this, Noonan claims that the distinction between Lumpl and Goliath are only apparent and their distinction does not render false the fact that Lumpl and Goliath refer to the same statue. Moreover, acceptance of Abelardian predicates is not merely an option available to the defenders of the constitution is identity thesis; it is the only option available to them.

Lynn Rudder Baker's argument against the notion of Abelardian predicates in her article 'Why Constitution is not Identity?' can be formulated as follows:

a. The meaning of the predicate 'might have been squeezed into a ball and not destroyed' depends on the meaning of the subject term (Goliath or Lumpl) if and only if the meaning of the predicate expressing persistence conditions depends on the meaning of the subject term to which it is attached.
b. It is false that the meaning of a predicate expressing persistence conditions depends on the meaning of the subject term to which it is attached.

---

c. It is false that the meaning of the predicate of the form 'might have been squeezed into a ball and not destroyed' depends on the meaning of the subject term (Goliath or Lumpl) to which it is attached.

The strength of this argument depends very much on the truth of premise (b). According to Baker, premise (b) is true because Lumpl/Goliath cannot have one set of persistence conditions relative to being Goliath (a statue) and another set of persistence conditions relative to being Lumpl (a piece of clay).

There seems to be two difficulties with Abelardian predicates here. First, either the predicate 'might have been squeezed into a ball and not destroyed' is true of the 'statue = piece of clay' or it is not. To claim that the predicate is both true and false of the 'statue = piece of clay' would be an obvious contradiction, for it amounts to saying that the 'statue = piece of clay' squeezed into a ball is both destroyed and not destroyed. So if the predicate is true, then the 'statue = piece of clay' cannot be Goliath and if the predicate is false, then the 'statue = piece of clay' cannot be Lumpl. Second, an Abelardian predicate is defined as a predicate whose reference can be affected by the subject term to which it is attached. One can doubt whether the reference of the predicate, 'might have been squeezed into a ball and not destroyed', can mean anything else if the subject term to which it is attached changes. It is difficult to conceive how this predicate can mean one thing when it is attached to Lumpl and another when attached to Goliath. Moreover, what else can it mean when it is attached to Goliath?

Constitution is Not Identity Theses

Co-location

Co-location is the most direct view under the 'constitution is not identity theses'. It holds that if Lumpl constitutes but is not identical to Goliath, then Lumpl and Goliath are simply spatially coincident. And it accepts the implausible consequence of having two objects in the same place at the same time without being related in any other away. For this reason, I reject the notion of co-location as a plausible solution to the problem of material constitution. The co-location theorists might object by saying that there is nothing implausible about having two objects in the same place at the same time. To this objection, I raise two questions. First, do Lumpl and Goliath share the same atoms? Yes and because they do, the atoms of Lumpl cannot overlap with those of Goliath and they count as one thing. But this is not necessarily an identity relation and I shall explain why later. Second, can the same atoms constitute two objects in the same place at the same time? No, unless one (constituting) object is part of another (constituted) object. The 'constituted object' refers to the 'constituting parts' and their structural configuration and relational properties. Note that the 'constituting parts' are distinct from the structural configuration and relational properties of the 'constituted object', even when the atoms of the 'constituting parts' and the 'constituted object' share exactly the same atoms without overlapping. I shall elaborate on this distinction later. To make things clear, I shall call 'constituting object' the material parts and call 'constituted object' the material object. Only a material object and its material parts that share the same atoms can exist in the same place at the same time, and no two material objects can be in the same place at the

same time. If this is true, then there is something implausible about having two material objects in the same place at the same time.[3]

Accidental Sameness

Michael Rea, in his article '*Sameness without Identity: an Aristotelian Solution to the Problem of Material Constitution*', uses Aristotle's notion of 'accidental sameness' to resolve the problem of material constitution. According to Aristotle, something called 'seated-Socrates' comes into existence as an 'accidental unity' when Socrates sits down. Though not a substance, it has a hylomorphic structure: Socrates refers to the matter and his seatedness refers to the form or the unifying principle. It comes into existence when seated Socrates is seated and goes out of existence when Socrates is no longer seated. The relation of accidental sameness obtains between Socrates and seatedness for only as long as they co-exist. Accidental sameness is not identity but a kind of numerical sameness. This relation is weaker than strict identity but stronger than co-location. Necessarily, if x and y are numerically the same at a certain time, then x and y share all of their parts in common at that time. Accidental sameness entails complete overlap of parts but co-location does not. For example, co-location allows an event and a material object to fully occupy the same space at the same without complete overlap of parts. Accidental sameness occurs whenever some matter is organised in several different ways

[3] One may argue that there is nothing implausible in having two objects of a different kind, comprised of the same atoms, in the same place at the same time. I agree on one condition, that the object of one kind must itself be a part of the object of another kind. For example, the atomic, chemical and biological components (kind) of a person x can be in the same place at the same time, because the atomic component (kind) of person x is part of the chemical component (kind) of person x, and the chemical component (kind) of person x is part of the biological component (kind) of person x.

at once, and whenever it occurs we can identify different kinds of objects in the same place but counted as one thing. It affords us an easy solution to the problem of material constitution. Whenever two objects x and y share all of the same parts but are related to their parts in different ways, x and y are numerically the same but not identical.

Rea states that Aristotle's notion of accidental sameness has the following facts: (1) it is the relation that holds between an accidental unity and its parent substance; (2) it is neither necessary identity nor contingent identity (if x and y are accidentally the same, they are in some way the same and in some way different; but if x and y are necessarily or contingently identical, there is no way in which they are different); and (3) it is a version of numerical sameness (if x and y are accidentally the same, then the are 'one in number' even if they are not 'one in being'). Elaborating on the fact (3), Rea says that we should count only one object in every region that filled by matter unified in some object-constituting way. We count one Goliath in every region that is filled by matter arranged Goliath-wise; we count one Lumpl in every region that is filled by matter arranged lumpl-wise; and we count one object in every region that is filled by matter arranged in either or both of these object-constituting ways. Thus, when we recognise Goliath and Lumpl in a particular region and deny that Goliath is identical with Lumpl, we are committed to the claim that there is matter in a particular region arranged both Goliath-wise and Lumpl-wise, and so being Goliath is something different from being Lumpl; but this is all consistent with there being one object in a particular region.

There are difficulties with this view. Aristotle tells us that accidental unity and its parent substance are 'one in number' but not 'one in being' at the same time. One can question how two or more beings existing in the same place at the same time can count as one entity. Aristotle's example says that a 'seated-Socrates' comes into existence on top of the Socrates-proper when seated and goes out of existence when he is no longer sitting. It seems problematic that there can be a indefinite number of Socrates' accidental unities, perhaps one for 'standing', one for 'walking', one for 'sleeping' etc to be treated as many different beings but counted as one thing. A related difficulty is that how can matter be arranged in more than one way (both Lumpl-wise and Goliath-wise) in the same place at the same time? And given an indefinite number of accidental unities of Socrates, which one counts as Socrates-proper (parent substance, primary being) in the first place? A more plausible alternative would be to treat all the accidental unities as different states (rather than different beings coming in and out of existence) of one material object constituted by material parts. The material object is capable of being in different states as a result of the material parts being arranged in different ways, thereby making the material object qualitatively distinct from its constitutive parts. I would elaborate on this view later in this chapter.

Baker's Essentialist Argument against 'Constitution is Identity' Thesis

In her article 'Why Constitution is not Identity?' Lynn Rudder Baker put forward an essentialist argument against the 'Constitution is Identity' thesis. Her argument has the following form:

a. x is essentially an F.

b. y is not essentially an F.

---

c. $x \neq y$

Applying this to Gibbard's example, we have

d. Goliath is essentially a statue.

e. Lumpl (a piece of clay) is not essentially a statue.

---

f. Goliath ≠ Lumpl

Premise (d) affirms that anything that existed and was not a statue (at all times of its existence) would not be a statue. It holds that being a statue is a property that a statue cannot lose without going out of existence, just as being three-sided is a property a triangle cannot lose without going out of existence. And it entails that if a particular statue (Goliath) ceases to be a paradigm statue (sortal term), then it goes out of existence. Hence, being a statue is essential to (or is definitive of) Goliath. Premise (e) affirms that it is possible that Lumpl (a piece of clay that constitutes Goliath) exists and is not a statue. A piece of clay can be used to construct many things other than a statue. Hence, being a statue is not essential to (or is not definitive of) Lumpl. Taken together, premises (d) and (e) states that Goliath has an essential property that Lumpl lacks. And inserting the hidden premise in the form of Leibniz's criterion for identity, we have conclusion (f): Goliath ≠ Lumpl. That is, (d) and (e) entail (f). The argument seems valid.

Baker then discusses an essentialist argument for the 'Constitution is Identity' thesis as a criticism against her essentialist argument against the 'Constitution is Identity' thesis:

g. x is essentially an F

h. x = y

---

i. y is essentially an F

Applying this Gibbard's example, we have

j. Goliath is essentially a statue

k. Goliath = Lumpl

---

l. Lumpl is essentially a statue

My own criticism against constitution as relative identity is that (h) is assumed to be true and is used as a premise to establish (i), or that (k) is assumed to be true and is used as a premise to establish (l). The problem of material constitution as I see it is to show whether or not (h) and (k) are true, and not whether or not (i) and (l) are true. In other words, the essentialist argument for the 'Constitution is Identity' thesis begs the question against the problem of material constitution as I see it. Baker holds that the opposite conclusion that Lumpl is not essentially a statue is stronger the premise Lumpl = Goliath because of a modal intuition that Lumpl could have existed without being a statue. Again, a piece of clay can be used to construct many things other than a statue. This argument is

similar to the one offered by the constitution as relative identity view above and Baker's criticism here is applicable against it. If Baker's constitution without identity view is correct, then Lumpl is not identical with Goliath because they possess a different essential property and persistence condition. But how would Baker then account for two distinct objects existing in the same place at the same time? In what follows, I would present Baker' constitution without identity view and assess its success in answering this question.

Baker's 'Constitution without Identity' Thesis and the Problem of Material Constitution

In her article '*Unity without Identity: A New Look at Material Constitution*', Baker states the problem of material constitution in the following way: For any x and y that are related as the piece of clay is to the statue that it makes up, either x is identical to y, or x and y are separate entities, independent of each other. She defines her position in between identity and separation. She rejects identity by stating that if x constitutes y, then $x \neq y$. This is because x and y could have different essential properties and persistence conditions. She rejects separation by holding that if x constitutes y, then they share many properties are spatially coincident. Hence, they are too closely related to be separate.

Let's elaborate this view using Gibbard's example. For Baker, identity is a necessary relation between objects. That is, if $x = y$, then necessarily $x = y$, and so x cannot differ from y in any respect, including respects in which x might have been, or might become, different from the way this is now. In other words, x and y share their modal properties –

which are properties of being possibly such and such, or of necessarily being such and such. If Lumpl and Goliath were identical, then by Leibniz's criterion for identity, there would be no property borne by Lumpl but not borne by Goliath, and no property borne by Goliath and not borne by Lumpl. But Lumpl could have existed without being a statue but Goliath could not. And Goliath has the aesthetic qualities essential to a statue but Lumpl has not. There is then a property borne by Lumpl but not Goliath, and a property borne by Goliath but not Lumpl. Therefore, the relationship between Lumpl and Goliath cannot be defined as identity. (See Baker's Essentialist Argument discussed above.)

The alternative seems to be one of separation (co-location), that Lumpl and Goliath are separate and independent individuals located in the same place at the same time. Baker disagrees and argues that Lumpl and Goliath are not just two independent individuals for two reasons. First, many of Goliath's aesthetic properties depend on Lumpl's physical properties – its weight distribution and shape. Second, Lumpl and Goliath are spatially coincident. Not only are they located in the same place at the same time, they have the same size, weight, shape, colour, texture and so on. Their similarity is no accident (hence rejects accidental sameness discussed above), for Goliath does not exists separately from Lumpl. Baker further argues that neither Goliath has Lumpl as a proper part nor Goliath is Lumpl plus something else. Hence, Lumpl is neither identical to nor separate from Goliath. Rather, the relation between Lumpl and Goliath is one of constitution.

Baker's constitution thesis has three important assumptions. First, constitution is a relation between two individual objects. Neither stuff (component parts of objects) nor

properties (features possessed by objects) are relata of the constitution relation. The statue Goliath (an object) and a piece of clay Lumpl (another object) figure as the relata in a constitution relation. The relation between Goliath (an object) and its clay parts (stuff), or the relation between Lumpl (an object) and its shape (properties) are not one of constitution. Here, it is quite clear that Baker rejects mereology – the thesis that the parts constitute the whole. Second, constitution is a contingent relation: Lumpl could have existed yet failed to constitute anything at all. If x constitutes y at some time, then the existence of x at that time does not by itself entail the existence of y. Third, many of the relational properties (e.g. aesthetic appeal) that make something (e.g. Goliath) the thing that it is (an art piece) are intentional. Baker defines an intentional property as follows: a property H is intentional if and only if H could not be exemplified in a world where no one ever one had a belief, desire, or any propositional attitude. Baker departs from tradition in taking certain relational and intentional properties to be essential properties. For example, Lumpl could exist in the absence of an art world but Goliath could not. In other words, it is essential for Goliath to be an art piece but not Lumpl. It is important also to recognise that not all constituted things are intentional.

Baker's constitution thesis has three following ideas – 'primary kinds', 'circumstances' and 'borrowed properties'. First, Baker requires the idea of primary kinds because constitution is a relation between two individual objects, and each individual object is fundamentally a member of exactly one kind, called a primary kind. To cite the question 'what is fundamentally or essentially x?' we cite x's primary kind by using a subject noun, for example 'a piece of clay' or 'a statue'. A primary kind is a kind of thing and

not just stuff (component parts of things) or properties (features possessed by things). An important feature of a primary kind is that an individual object ceases to exist if its primary kind (essential) property ceases to exist. Second, Baker requires an idea of circumstances to answer the question 'in virtue of what is x the kind of thing that it is?' Many properties can be instantiated only in certain circumstances. Baker formulates this idea as follows: where F (Lumpl) and G (Goliath) are distinct primary kind properties, it is possible that F (Lumpl) exists without any spatially coincidental G (Goliath). However, if an F (Lumpl) is in G-favourable circumstances, then there is a new entity, a G (Goliath), that is spatially identical to F (Lumpl) but not identical to it. Third, Baker requires the idea of borrowed properties to answer the following question: suppose that object x has property H, in virtue of what does x has H? Baker's answer is that x has H in virtue of constituting something that has H or being constituted by something that has H. An important feature of constitution requires a distinction between properties that are borrowed and properties that are not borrowed. Baker formulates the idea of borrowed property as follows: Suppose objects x (Lumpl) and y (Goliath) are constitutionally related, H (aesthetic appeal) is a borrowed property of x (Lumpl) at t if x (Lumpl) having H (aesthetic appeal) at t derives exclusively from x's (Lumpl's) being constitutionally related at t to y (Goliath) that has H (aesthetic appeal) independently at t. The idea of borrowed property is essential to Baker's account because x (lumpl) and y (Goliath) are seen as distinct objects even when they are related constitutionally.

In her constitution without identity thesis, Baker succeeded in explaining the qualitative differences (essential properties / persistence conditions) between two constitutionally

related items and in establishing the intimate relation between them. But in conceiving constitutionally related items as distinct objects (primary kinds) borrowing properties from each other, the problem of there being two objects occupying the same space at the same time remained. To address this problem, Baker's critics have urged her to define constitution as a mereological relation, which she rejected. I would want to explore the idea of constitution as a mereological relation without identity next and assess its success as a solution to the problem of material constitution.

Constitution is Mereological Relation without Identity Thesis

Merelogical Relation and Identity Statements

As indicated earlier, my focus here is to specify the relationship between the material parts and the material object they constitute, or between xs and the fusion of xs. The constitutive relation between xs and the fusion of xs is best construed as a mereological relation without identity. I believe that this is the most plausible solution to the problem of material constitution, where qualitative differences can be adequately accounted for without admitting that it is possible for two objects to occupy the same space at the same time. I would like to formulate a notion of mereological relation without identity by defining it against the various statements of identity relation between the material object and the material parts that constitute it as follows:

1. Xs and the fusion of xs occupy the same space at the same time.

2. What affects xs affects the fusion of xs and vice versa.
3. Xs and the fusion of xs share the same parts.
4. Xs and the fusion of xs count as one thing.
5. The fusion of xs exists when xs exist.
6. The fusion of xs is nothing more than the xs.
7. The fusion of xs is not distinct from the xs.
8. The fusion of xs is fully described when xs are described.

To accept all these points is to accept strict identity between xs and the fusion of xs as spelt out by Lewis above. For those who accept Leibniz's criterion of identity, it is not possible to accept (1) to (4) and reject (5) to (8) at the same time. They hold that the truth of points (1) to (4) entails the truth of points (5) to (8). But in accepting (5) to (8), strict identity cannot account for the possibility that xs and the fusion of xs can have at least one differing property (qualitative distinction). I am inclined to accept points (1) to (4) without qualifications and reject points (5) to (8) outright. To do this, I have to make an important distinction between the material parts that make up the material object and ways the material object are. John Heil made this distinction in his book '*From an Ontological Point of View*'. I hold that xs refer only to the material parts that make up the material object but the fusion of xs includes both the material parts that make up the material object as well as the ways the material objects are. It is my acceptance of this mereological relation (xs are parts of the fusion of xs) that commits me to the points (1) to (4) and it is my denial of qualitative identity that leads me to reject points (5) to (8). In what follows, I will first discuss how the remaining notions of identity would treat

these eight points and then elaborate on my treatment of these eight points which defines my solution to the problem of material constitution.

Identity Statements and the Constitution as Identity Theses

The notion of contingent identity would accept points (1) to (8) in describing the relation between xs and the fusion of xs, but not necessarily. It is possible that any or all points from (1) to (8) do not describe the relation between xs and fusion of xs. That is, it is just incidental that xs and the fusion of xs are identical but it remains possible that xs and the fusion of xs are not identical. If points (1) to (8) apply to the relationship between xs and the fusion of xs and they are hence identical, then it is not possible for them to be distinct objects. Rejecting the identity relation would mean that it is possible for two distinct objects to occupy the same place at the same time, which is highly implausible. Since contingent identity accepts that it is possible for xs and fusion of xs to be distinct objects, and thereby accepts the possibility of having two distinct objects occupying the same place at the same time, I reject it. In my view, the mereological relation between xs and the fusion of xs (that xs are parts of the fusion of xs) implies that they are not two objects occupying the same space at the same time. And the mereological relation allows for the possibility of qualitative distinction between xs and the fusion of xs.

The notion of relative identity would define a sortal term F and then determine whether xs is intrinsically exactly like F and whether the fusion of xs is intrinsically exactly like F. If xs is intrinsically exactly like F and the fusion of xs is intrinsically exactly like F,

then both xs and the fusion of xs are identical in relation to sortal term F. At most, this tells us that xs and the fusion of xs belong to the same sort, class or type of things. It does not tell us whether points (1) and (8) describe the relation between xs and the fusion of xs, and stops short of telling us whether or not xs and the fusion of xs are the same object. And if relative identity cannot establish whether or not xs and the fusion of xs are the same object, then it has failed to solve the problem of material constitution as I see it.

The Abelardian predicates thesis seems to accept points (1) to (6) as xs and the fusions of xs refer to the same object, even when there are two different senses in talking about them. Abelardian predicates seem to be incompatible with (7) and (8) because xs and the fusion of xs can have apparently distinct properties and hence predicates with different references depending on how the subject term is defined. Hence, xs and the fusion of xs can be described in different ways even when they are identical. But if Baker is correct, then it is possible for xs and fusion of xs (different senses of referring to one object) to have different persistence conditions (different senses of referring to one predicate) according to Abelardian predicates. If it is possible for xs and the fusion of xs to have different persistence conditions, then it is possible for one to exist whiles the other ceases to exist. And if it is possible for one to exist while the other ceases to exist, then they cannot be the same object. Baker's counter argument shows that xs and the fusion of xs have concrete differences – their persistence conditions. So, the notion of Abelardian predicates cannot show whether or not xs and the fusion of xs are the very same object. Hence, I reject the notion of Abelardian predicates solving the problem of material constitution.

Identity Statements and the Constitution Without Identity Theses

The notion of accidental sameness is construed as a kind of numerical sameness. This relation is weaker than strict identity but stronger than co-location. It seems to accept points (1) to (4) and affirms numerical sameness. Accidental sameness occurs whenever some matter is organised in several different ways at once, and whenever it occurs we can identify different kinds of objects in the same place but counted as one thing. Whenever xs and the fusion of xs share all of the same parts but are related to their parts in different ways, x and y are numerically the same but not identical. And whenever we recognise xs and the fusion of xs in a particular region and deny that xs is identical with the fusion of xs, we are committed to the claim that there is matter in the region arranged both in the ways of xs and the fusion of xs, and being xs is something different from being the fusion of xs. My questions here are twofold. First, if xs and the fusion of xs share all of the same parts, how can they be different objects? Second, if xs and the fusions of xs are different kinds of objects (even if some are accidental) in the same place at the same time, how can they be the same numerically? In holding that xs and the fusion of xs share all the same parts, and that there can only be one object in the same place at the same time necessarily, I reject the notion of accidental sameness. Instead, I believe that mereological relation fulfils these two conditions better.

Baker's position accepts points (1) and (2) and denies the rest. Her rejection of (3) and (4) is based on the assumption that the relatas of the constitution relation are distinct

objects with different essential properties, in particular persistence conditions. Her rejection of (5) and (6) is based on the assumption that constitution is a contingent relation – it is possible for xs to exist without becoming the fusion of xs, and the fusion of xs only come into existence when xs are in certain circumstances. Her rejection of (7) and (8) can be explained by the assumption that fusion of xs can have certain relational or intentional properties which xs lacked. As a result, her position is that xs and the fusion of xs are constitutionally related without being identical. Although Baker insists that xs and the fusion of xs are constitutionally related, she accepts the implausible view that xs and the fusion of xs are two distinct objects occupying the same space at the same time. This implausible view can be avoided by holding that xs and the fusion of xs are mereologically related without being identical.

Identity Statements and Constitution as Mereological Relation without Identity

My position is that constitution is a mereological relation without identity. Affirming mereological relation by accepting points (1) to (4) is essential to avoid the implausible position that there are two objects occupying the same space at the same time. To establish mereological relation, I need points (1) and (3) to arrive at point (4), and later derive point (2) from point (4). If xs and the fusion of xs occupy the same space at the same time and if xs and the fusion of xs share the same parts, then xs and the fusion of xs count as one thing. And if xs and the fusion of xs count as one thing, then what affects xs affects the fusion of xs and vice versa. My view differs from Baker's in affirming (3) – the mereological relation between xs and the fusion of xs; and affirming (4) – the denial

of xs and the fusion of xs as different but co-located and constitutionally related objects. I believe that the relatas of any constitution relation are not distinct objects but rather material parts and material objects (= fusion of material parts = material parts configured in certain ways under certain circumstances possessing certain relational properties – dispositions and qualities). That is, material object = material parts + ways the materials parts are; and ways material parts are = structural configurations of material parts + relational properties (dispositions and qualities). By structural configuration, I mean the arrangement and organisation of material parts when conceived statically, as well as the connections and interactions between material parts when conceived dynamically. When material parts are structurally configured in certain ways under certain circumstances, it possesses certain relational properties (at the level of material object) not possessed by its material parts.

Having drawn this distinction between xs (material parts) and the fusion of xs (material objects), I am in a position to reject points (5) to (8).[4] Rejecting qualitative identity by rejecting points (5) to (8) is essential to account for the qualitative differences between xs (material parts) and the fusion of xs (material objects). I reject point (5) because the material parts can still exist without constituting the material object. It is possible that the material parts lack the fusion or the structural configuration under certain circumstances to be the material object. Point (6) is rejected for the similar reason because the material

[4] One may argue that the simultaneous acceptance of (1) and rejection of (5) is inconsistent. I disagree. Acceptance of (1) affirms that xs and 'fusion of xs' occupy the same place at the same time as they share exactly the same atoms. While both xs and 'fusion of xs' include the exact same atomic parts, xs excludes the configuration of these atomic parts and their relational properties but 'fusion of xs' includes them. This is consistent with the rejection of (5), which denies that 'fusion of xs' exists when xs exists, because it is possible for the atomic parts (xs) to exist without the configuration of these atomic parts and their relational properties (all of which are necessary for 'fusions of xs' to exist) at the same time.

parts may lack the fusion or structural configuration under certain circumstances to be the material object. But rejection of points (5) and (6) are compatible with the mereological relation between xs and the fusion of xs. The whole (material object) is simply ways the material parts are and there is qualitative difference between material parts and ways the material parts are. Point (7) is rejected because the material object (= fusion of material parts = material parts configured in certain ways under certain circumstances possessing certain relational properties – dispositions and qualities) encompasses more than its material parts and hence is distinct from them. Point (8) is rejected for the similar reason – exactly because material objects encompass more than its material parts, material objects cannot be fully described by its parts. Again, rejection of points (7) and (8) are compatible with the mereological relation, for the material parts are the constituents of the material object.

The mereological relation I hold differs significantly from the widely held constitution as identity view. According to the identity view, parts that constitute the whole are seen as identical to the whole. Let parts (xs) be x, whole (fusion of xs) be y, and let y be part of a greater whole z. Read this way, the mereological relation between x, y and z are:

a. Reflexive – x is identical to itself.
b. Symmetric – if x is identical to y, then y is identical to x.
c. Transitive – if x is identical to y, and y is identical to z, then x is identical to z.

According to my constitution without identity view, parts that constitute the whole are not seen as identical to the whole. Read this way, the mereological relation between x, y and z are:

d. Irreflexive – x does not constitute itself.
e. Asymmetric – if x constitutes y, then y does not constitute x.
f. Transitive – if x constitutes y, and y constitutes z, then x constitutes z.

Baker held a similar view that the constitution relation is irreflexive, asymmetric and transitive but she does not conceive the constitution relation as a mereological relation. My view of mereological relation has another feature:

g. Non-causal – if x constitutes y, then x does not cause y (in the efficient sense).

Applying my position to Gibbard's example, Lumpl refers to the material parts (xs) and Goliath refers to the material object (fusion of xs). Acceptance of points (1) to (4) makes Lumpl a part of (mereologically related to) Goliath. Denial of points (5) to (8) makes Lumpl qualitatively distinct from Goliath, because Goliath (a statue) refers to Lumpl (a piece of clay) in a specific configuration under circumstances with specific dispositions and qualities. Lumpl does not constitute itself. Rather, Lumpl is constituted by atoms (components other than itself) and is hence irreflexive. Lumpl constitutes Goliath when it has certain configuration, dispositions and qualities but Goliath (that includes structural configuration and relational properties) does not constitute Lumpl (that does not include

structural configuration and relational properties).[5] Hence, the relationship between Lumpl and Goliath is asymmetric. And if Lumpl constitutes Goliath, then Lumpl does not (efficiently) cause Goliath. Hence, the relationship between Lumpl and Goliath are non-causal.

My position can be contrasted with the key notions of Baker's view. First, I refrain from using the term 'primary kinds' as specified by Baker, where they refer to kinds of things and cannot be parts or properties.[6] Concepts of material parts, structural configuration and relational properties (dispositions and qualities) figure significantly in my view of constitution. Baker denies this view because she rejects mereology. Second, I believe that Lumpl and Goliath are not two distinct, co-located, and constitutionally related primary kinds borrowing properties from one another. Rather, the 'constituted' material object (Goliath) is made up of the 'constituting' material parts (Lumpl) structurally configured in certain ways with certain relational properties (dispositions and qualities), located in one particular space at a particular time. Third, I do agree with Baker's notion of circumstances – that certain environmental situations can be conditions for material parts taking on certain structural configurations. Fourth, my view is compatible with Baker's view that Lumpl and Goliath maybe qualitatively different and may even have different persistence conditions. Lumpl that loses its structural configuration and

[5] Lumpl and Goliath are constituted by exactly the same atoms and hence occupy the same space at the same time. The distinction between Lumpl and Goliath lies in that 'Goliath' refers to the atomic parts, the ways they are arranged, as well as the possession of relational properties when the atomic parts are so arranged, while 'Lumpl' refers only the atomic parts. Unlike atoms, configuration and relational properties (conceived as dispositions and qualities) do not occupy space but this does not mean that they do not exist.

[6] I believe that primary kinds can both constitute other primary kinds or be constituted by other primary kinds. Constituting primary kinds refer to parts and constituted primary kinds refer to wholes (or objects). But I refrain from using the term primary kinds in my essay to avoid the misconception that primary kinds always refer to independent objects (or wholes) and not parts.

relational properties but retains its material parts may still persist as a piece of clay even if it ceases to be Goliath. Conversely, Goliath may not persist when Lumpl ceases to exist because Lumpl is the material part of Goliath. When the material parts ceases to exist, its structural configuration and relational properties would cease to exist as well.

Let us now return to Lewis' question raised in the first section of this paper: is mereology ontologically innocent? If we accept the constitution is identity thesis, that mereological relation is reflexive and symmetric, then it is certain that mereology is innocent. Baker's constitution without identity thesis rejects mereology precisely because of its innocence. On my constitution as mereological relation but not identity view, where mereological relation is both irreflexive and asymmetric, the answer is both yes and no. On the one hand, the answer is yes because:

a. If fusion of xs (Xf) = aggregates of x parts (X1, X2, X3), then Xf is composed of no other non-overlapping material parts in addition to X1, X2, and X3.
b. If fusion of xs (Xf) = aggregates of x parts (X1, X2, X3), then Xf decomposes exhaustively into non-overlapping material parts X1, X2 and X3.

On the other hand, the answer is no because Xf does not refer only to its material parts X1, X2, and X3, but also to the structural configuration of the material parts X1, X2 and X3 that possess certain relational properties. For an adequate account of ontology, we must take into account the material parts, the structural configuration, as well as the relational properties of material objects. The structural configuration and relational

properties of Xf do not add additional material (or nonmaterial) parts to Xf, they are just 'ways Xf is'. In their article 'The Ontological Turn', C.B. Martin and John Heil put it in the following way:

"Objects can have parts, but an object's properties are not its parts, they are the particular ways the object is." (C.B. Martin and John Heil 1999, P.45)

As the material object refers to more than just its material parts, mereology is not as innocent as Lewis wants it to be.

Mereological Essentialism, Mereological Nihilism and Merelogical Supervenience

At this juncture, I would like to explore how my position relates to the other mereological theses: mereological nihilism (MN), mereological essentialism (ME) and mereological supervenience (MS). In particular, I would like to reject MN by insisting that mereology is crucial for an adequate account of ontology, reject ME by insisting that they are not necessarily entailed by mereology, and accept a non-identical or non-reductive notion of MS.

Mereological Nihilism (MN)

MN is the view that objects with parts do not exist and only basic building blocks without parts exist. The world only contains stuffs or masses of matter and they come in different

quantities. Our standard way of quantifying is at most a way for the mind to project individual object and part distinctions onto the world and it is our misconception to see the world as full of objects with parts. MN amounts to a denial of mereology. Peter Unger argues for this position in his paper '*There are No Ordinary Things*'. In my view, the world consists of material objects and the ways they are. The 'ways material objects are' are accounted for by their structural configurations and relational properties (dispositions and qualities). The world does not consist of stuffs or masses of matter simply categorised into objects and parts arbitrarily by our minds. Rather, stuffs or masses of matter are organised into myriad individual structural configurations with certain relational properties (dispositions and qualities); and it is material parts that figure in the structural configurations of material objects.[7] As mentioned earlier, an adequate account of ontology must include the material parts, their structural configurations, as well as the relational properties (dispositions and qualities) of material objects. In denying mereology, MN is unable to offer an adequate account of ontology. Thus, mereology is required for a plausible account of constitution.

Mereological Essentialism (ME)

ME is the view that objects have their parts essentially – that is, if an object would lose (or gain) a part, it would cease to exist. Roderick Chisholm argues for ME in his paper

[7] Some philosophers who emphasises Kant's transcendental idealism thesis may argue that it is our mind that structure reality according to the categories and hence perceptual and conceptual distinctions are made in our mental representations of reality. Without the mind and its categories to make such distinctions, reality could just be stuffs or masses of matter. I disagree because I emphasises Kant's objectivity thesis – that reality consists of objects (and not just stuffs or masses of matter) existing independently of perceivers, even though we represent them according to the categories. It is worth noting that according to Kant, both theses are necessary conditions for experience.

'*Parts as Essential to their Wholes*'. The idea is this: rejection of ME means that objects could survive the loss and replacement of its parts but we would not know how much loss or replacement of parts objects could take before it would cease to exist. And the best way to avoid this problem would be to accept ME. One objection raised is that if ME is true, then no objects could survive change. This makes change in objects impossible. ME holds that objects cease to exist whenever they lose (or gain) an atom and new related objects constantly arise to take their place, thus introducing a new problem of the many. If it is true that objects change by losing and gaining parts and properties, then ME is false. Also, rejecting ME avoids the counterintuitive view that objects constantly cease to exist and replaced by other related objects whenever they lose (or gain) an atom. The problem of rejecting ME arises from our definition or conception an object as a rigid or fixed set of parts or atoms rather than the nature of objects themselves. And I advocate a more flexible and fluid way of defining and conceiving objects as well as their parts and properties – that it is possible for an object to loose or gain some parts or atoms (change or be qualitatively different) over time. While admitting that objects cease to exist when they lose significant parts, configurations and properties, I would assert that they survive the loss of certain parts, configurations and properties. There is no question of objects surviving the loss of some of their parts. Alvin Plantinga, in his paper '*On Mereological Essentialism*', provides an excellent example to this flexible and fluid view. As organisms, human beings survive by having our parts replaced by metabolism (cell replacement), tissue/organ transplants, or even cutting and regeneration of our fingernails and hair. All these processes do not seem to lead to the cessation of the human being's existence. Perhaps a person could not survive the loss of their brain. It is worth noting

that mereological relation is retained. Biological systems, organs, tissue and cells remain part of the human organism despite loss and replacement of certain parts. Hence, mereology does not necessarily entail ME.

Mereological Supervenience (MS)

In his '*Physicalism and Panexperientialism: Response to David Ray Griffin*', Jaegwon Kim says of MS:

"Mereological supervenience only asserts that the properties of the whole are determined, or fixed, by the properties and relations that characterise its parts. That only means that if two wholes are microstructurally identical, they must exhibit the same macroproperties – and the same causal powers. And these can be new causal powers; it is only that they are determined by microstructure. Mereological supervenience does not say, or imply, that the properties of the whole are identical with properties of their microconstituents. As emergentists too would say: such properties as inflammability, ductility, and temperature of macro-objects are not among the properties of individual molecules or atoms. For both emergentists and (most) physicalists, they are genuine properties and causal powers, which supervene on, or are determined by the microstructural make-up of the objects that have them. Emergentists and physicalists would stress that the structural configuration, no less than the intrinsic properties of microconstituents, is crucial determining what macro-properties are exhibited by a whole. For wholes – anyway, those of interest to us – are structures, not mere

assemblages of atoms and particles, and the very same atoms and particles configured in different structural relationships can, and do, exhibit very different properties and causal powers at the level of wholes. And many of these properties are not had by the wholes' microconstituents. This is completely consistent with physicalism and in particular mereological supervenience and micro-reductionism. Thus two points: first, determinism must be sharply distinguished from identity, and second, structure is crucially important." (Jaegwon Kim 1999, P.33-34)

According to Kim, MS asserts that the properties of the whole are determined by, fixed by, or dependent on and at the same time not possessed by, identical with, or reducible to the properties of its parts or their structural configuration. In causal terms, MS asserts that the causal powers of the macroproperties of the whole are determined by, fixed by, or dependent on and at the same time not possessed by, identical with, or reducible to the causal powers of microproperties of its parts or their structural configuration. When the microproperties of parts possess certain structural configuration, the whole possesses new and genuine macroproperties. In this way, MS does not imply a causal relation between the whole and its parts. This is because if 'the microproperties of the parts and structural configuration' and 'the whole and its macroproperties' are instantiated at the same time, then both 'the microproperties of the parts and structural configuration' and 'the whole and its macroproperties' do not figure as the antecedent of (prior to) or the consequent of (posterior to) one another in a conditional (causal) relation. So, the casual powers of the whole and its macroproperties are conserved even when they are determined or fixed by

the microproperties of the parts and their structural configuration. Kim makes this point clear in his *'Mind in a Physical World'* with examples:

"This table has a mass of ten kilograms, and this property, that of having a mass of ten kilograms, represents a well-defined set of causal powers. But no micro-constituent of this table, none of its proper parts, has this property of the causal powers it represents. H2O molecules have causal powers that no oxygen or hydrogen atoms have. A neural assembly consisting of many thousands of neurons will have properties whose causal powers go beyond the causal powers of the properties of its constituent neurons, or subassemblies, and human beings have causal powers that none of our individual organs have. Clearly then macro-properties can, and in general do, have their own causal powers, powers that go beyond the causal powers of their micro-constituents." (Jaegwon Kim 1998, P.85)

It is clear that Kim does not equate MS with mereological reduction, and he does not identify the whole with the sum of its parts. What MS amounts to is this: MS asserts that certain constituting material parts in certain structural configurations are necessary and sufficient conditions for the constituted whole and its relational properties, in such a way that when certain constituting material parts in certain structural configurations are instantiated, the constituted whole and its relational properties are instantiated. This non-identity and non-reductive notion of MS is compatible with the constitution without identity thesis.

Constitution and its Implication for the Mind-Body Problem

Baker's Constitution View of the Person and the Mind-Body Problem

In '*On Making Things Up: Constitution and Its Critics*', Lynn Rudder Baker refines her constitution thesis to construct a plausible view of the human person, to derive a plausible solution to the mind-body problem, and to defend her thesis against critics. The main idea of constitution is this: when a thing of one primary kind is in certain circumstances, a thing of another kind – a new thing with new causal powers – comes to exist. It describes the relationship between pieces of clay and statues, and DNA molecules and genes. When a certain combination of chemicals is in a certain environment, then a new thing – an organism – comes into existence. And when an organism evolves in certain ways, then a new thing – a person – comes into existence. But a statue is a different kind of thing from a piece of clay, a gene is a different kind of thing from DNA molecules, an organism is a different kind of thing from the combination of chemicals, and a person is a different kind of thing from the organism. Constitution makes an ontological difference. When a piece of clay constitutes a statue, it does not acquire the property of being a statue. Rather, a new kind with new causal powers and new persistence conditions – a statue – constituted by that piece of clay, comes into existence. Similarly, when an organism constitutes a person, it does not acquire the property of being a person. Instead, a person – a new kind with new causal powers and persistence conditions – comes into existence. Constitution relation is asymmetric. The piece of clay constitutes the statue and not vice versa. And the organism constitutes the person and not vice versa. This

asymmetry introduces some sort of ontological hierarchy. As one kind ascends the ontological ladder, it possesses greater causal powers and greater ontological significance. The asymmetry of constitution relation implies its non-reductionism.

Baker differentiates the constitution view from substance dualism in the following ways. First, substance dualism implies that there are two basic kinds of things but Baker's constitution view holds that there are countless basic kinds of things that cannot be reduced to two. Second, substance dualism conceives mind as an immaterial substance and body as a material substance but the constitution thesis holds that the mind is not a substance at all. Only the person and the body are substances, both of which are fully material, and neither of which are bearer of mental properties. Hence, the constitution view can be seen as a materialist account of the physical world. Third, substance dualism regards the relation between mind and body as causal but Baker holds that the relationship between person and body is non-causal. Baker then states three reasons to accept the constitution view of material objects. First, the constitution view is a comprehensive view of the natural world and it enables a plausible formulation of the materialism thesis: Materialism is true if and only if every concrete thing that exists at t is either a fundamental particle or is ultimately constituted by an aggregate of fundamental particles at t. So construed, materialism is compatible with the constitution view. Second, the constitution view is able to account for the different kinds of things in the world. It does not explain away the things that everyone believes exist and cannot avoid believing exist. So, a view that recognises the existence of many kinds is more natural than one that does not. Third, the constitution view is able to situate persons in the

natural word without detaching them from their animal nature. Both humans and animals are similar biological kinds.

Constitution as Mereological Relation without Identity and the Mind-Body Problem

For Baker, the idea of constitution is this: when a thing of one primary kind (a piece of clay) is in certain circumstances, a thing of another kind (a statue) – a new thing with new causal powers – comes to exist. Similarly in the person-body relation: when a certain combination of chemicals is in a certain environment, then a new thing – an organism with new causal powers and persistence conditions – comes into existence, and when an organism evolves in certain ways, then a new thing – a person with new causal powers and persistence conditions – comes into existence. Constitution makes an ontological difference. As a thing of one primary kind ascends the ontological ladder, it possesses greater causal powers and greater ontological significance. We can decipher a layered model of reality in Baker's position, where chemical kinds constitute biological kinds with greater causal powers (human bodies) under certain circumstances, and biological kinds constitute psychological kinds (persons) with greater causal powers under certain circumstances. So Baker's position has the following features:

a. When a simpler kind constitutes a more complex kind, there is more than one distinct object occupying the same space at the same time. So, the combinations of chemicals that constitute the organism and the organism that constitute the person are distinct objects occupying the same space at the same time.

b. When a simpler kind constitutes a more complex kind, the simpler kind does not acquire the property of being the complex kind. So, when the combinations of chemicals constitute an organism, the combination of chemicals does not acquire the property of being the organism, and when the organism constitutes the person, the organism does not acquire the property of being a person.

c. When a simpler kind constitutes a more complex kind, the simpler kind does not cause the complex kind to exist. So, when the combinations of chemicals constitute an organism, the combination of chemicals does not cause the organism to exist, and when the organism constitutes the person, the organism does not cause the person to exist.

d. When a simpler kind constitutes a more complex kind, the complex kind has more ontological significance and hence greater causal powers than the simpler kind. So, the combination of chemicals has less ontological significance and less causal power than the organism it constitutes, and the organism has less ontological significance and less causal power than the person it constitutes.[8]

My position rejects (a) and (b), accepts (c) outright and accepts (d) with qualifications. I reject (a) because a person refers to an organism configured in certain ways under certain suitable circumstances with certain psychological dispositions and qualities, and an organism refers to the combination of chemicals having certain configuration under certain suitable circumstances with certain biological dispositions and qualities. A person

[8] Baker did not specify what she means by complex kinds having more ontological significance and causal powers than simple kinds. Perhaps she means that an organism can perform higher functions that are denied to chemicals constituting it, and that a person can perform higher functions that are denied to organisms constituting him/her.

and its constitutionally related organism and combination of chemicals do not refer to separate and independent primary kinds or material objects. Rather, the combination of chemicals constitutes but is not identical to the organism, and the organism constitutes but is not identical to the person. Persons are just ways organisms arranged in certain manners with certain properties are, and organisms are just ways particular combinations of chemicals arranged in certain manners with certain properties are. Hence, it is not the case that there is more than one distinct object occupying the same space at the same time. I reject (b) because a person has certain psychological properties (dispositions & qualities) when the organism that constitutes it is configured in a certain way and an organism similarly has certain biological properties when the chemical components that constitute it are configured in certain ways. Hence, material parts constitute material objects by having certain properties when they are arranged in certain ways. I accept (c) because it is not the case that the organism causes the person. Rather, the causal power of the organism is part of (or constitutes) the causal power of the person. The causal power of the person comprises of the organism (biological parts), its configuration (as a human being), and its dispositions and qualities (psychological properties). But how is the person more ontologically significant and causally powerful than the organism that constitutes it? It is not that the person as a thing of a primary kind has more ontological significance and causally powerful than a thing of another primary kind, the organism, that constitutes it. Rather, I accept (d) because the person refers to its constituent biological parts together with its structural configuration and psychological properties (and hence encompasses more than just its biological parts). And the organism only

refers to its constituent chemical parts with its structural configuration and biological properties (and hence encompasses more than just its chemical parts).

The notion of constitution as mereological relation without identity and its implication for the mind-body problem can be clarified by the following illustration. Consider how the below items belonging to the same person are related to each other:

Sub-atomic Particles → Atoms → Molecules → Cells → Tissues → Organs → Systems → Person

The first thing to say about relationship of the items is this: items on the left constitute the items on their right. The constitution thesis I accept has the following characteristics, which can be expressed through the example

- Mereological relation – items on the left are parts of the items on their right. Assuming that the items constitute the same person, they are described by the acceptance of points (1) to (4) above. That is, (1) the constituting items on the left and the constituted items on the right occupy the same space at the same time; (2) what affects the constituting items on the left affects the constituted items on the right and vice versa; (3) the constituting items on the left and the constituted items on the right share the same parts, and (4) the constituting items on the left and the constituted items on the right count as one thing. It is not that items on the right are objects of certain primary kinds distinct from items on the left. Rather, the

constituting items on the left that make up the constituted items on the right are one object occupying the same space at the same time. Where items on the left refer to the constituting parts in certain configurations, items on the right refer to the constituted whole and their relational properties.

- Non-identity relation – items on the left are not identical to the items on their right. Assuming that the items constitute the same person, they are described by the rejection of points (5) to (8) above. That is, (5) items on the right exist when items on the left exist. (5) is false because it is possible for items on the left to exist without constituting items on their right; (6) items on the right is nothing more than items on the left. (6) is false because items on the right are constituted by items on their left when they are configured in certain ways with certain dispositions and qualities, and hence items on the right encompass more than items on their left; (7) items on the right are not distinct from the items on their right. (7) is false because of the same reason as (6); (8) items on the right are fully described by items on their left. (8) is false because items on the right encompass items on their left plus their configurations, dispositions and qualities. So, items on the left cannot be complete descriptions of items on their right.
- Non-causal relation - items on the left do not cause the items on their right (in the efficient sense). Rather, the causal power of the items on the left is part of (or constitutes) the causal power of the items on the right. It is not that items on the right have independent and distinct causal powers on their own, but that the causal powers of items on the right comprise of the items on their left as parts, their configuration and their dispositions and qualities.

In contrast with Baker, my idea of constitution is this: when parts (xs = piece of clay = Lumpl) have certain structural configuration under certain circumstances, a whole (fusion of xs = statue = Goliath) have certain relational properties (dispositions and qualities) that is not identical with or reducible to its material parts and structural configuration alone. The causal powers of the whole and its relational properties are determined by, fixed by or dependent on, but not possessed by, identical with, or reducible to its parts or structural configuration. And the whole and its relational properties are not caused by its parts or structural configuration. Similarly in the person-body relation: when a certain combination of chemicals with a certain structural configuration in a certain environment is instantiated, an organism with new dispositions and qualities is instantiated, and when an organism in a certain structural configuration in a certain environment is instantiated, a person with new dispositions and qualities is instantiated. I agree with Baker that constitution makes an ontological difference in another sense, that the constituted object is more than the sum of its non-overlapping constituting parts because its structural configuration and relational properties (ways objects are) do figure in ontology, even when the object decomposes exhaustively into its non-overlapping parts. In terms of the mind-body relation, the 'body' refers to material parts and their structural configuration, the 'mind' refers to the relational properties (dispositions and qualities), and the 'person' refers to the material object as a whole. Applying these notions back to Gibbard's example, 'body' refers to Lumpl, 'mind' refers to the aesthetic properties of a statue, and 'person' refers to Goliath. This completes my solution to the problem of material constitution and its implication for the mind-body problem.

Conclusion

In this chapter, I have first argued that the 'constitution is identity' solutions either offer inadequate accounts of qualitative differences or if they did, the notion of identity (as we understand it) is significantly compromised. Second, I have argued that the 'constitution is not identity' solution cannot avoid the possibility of having more than one object occupying the same place at the same time. Third, I have proposed the 'constitution as mereological relation without identity' solution both to offer an adequate account of qualitative differences and to avoid the possibility of having more than one object occupying the same place at the same time. Lastly, I have applied this solution as a plausible conceptual component of non-reductive physicalism to the mind-body problem.

Reference:

Baker, Lynne Rudder (1997), 'Why Constitution is Not Identity', *The Journal of Philosophy Vol.94 December*, P.599-621

Baker, Lynne Rudder (1999), 'Unity without Identity', *Midwest Studies in Philosophy, XXIII*, P.144-165

Baker, Lynne Rudder (2002), 'On Making Things Up: Constitution and its Critics', *Philosophical Topics Vol.30 No.1 Spring*, P.31-51

Chisholm, Roderick (1973), 'Parts as Essential to their Wholes', *Review of Metaphysics Vol.26*, P.581-603

Gibbard, Alan (1975), 'Contingent Identity', *Journal of Philosophical Logic Vol.4*, P.187-221

Heil, John (2003), *From an Ontological Point of View* (Oxford University Press)

Johnston, Mark (1992), 'Constitution is Not Identity', *Mind Vol.101*, P.90-105

Kim, Jaegwon (1998), *Mind in a Physical World* (MIT Press)

Kim, Jaegwon (1999), 'Physicalism and Panexperientialism: Response to David Ray Griffin', *Process Studies Vol.28*, P.28-34

Martin, C.B. and Heil, John (1999), 'The Ontological Turn', *Midwest Studies in Philosophy XXIII*, P.34-60

Noonan, Harold (1991), 'Indeterminate Identity, Contingent Identity and Abelardian Predicates', *Philosophical Quarterly, April*, P.183-193

Noonan, Harold (1993), 'Constitution is Identity', *Mind Vol.102*, P.133-146

Plantinga, Alvin (1975), 'On Mereological Essentialism', *Review of Metaphysics Vol.28*, P.468-76

Rea, Michael (1998), 'Sameness Without Identity: An Aristotelian Solution to the Problem of Material Constitution', *Ration (new series) XI 3December*, P.316-28

Unger, Peter (1979), 'There are No Ordinary Things', Synthese 41, P.117-54

Van Inwagen, Peter (1994), 'Composition as Identity', *Philosophical Perspectives 8 – Logic and Language*, P.207-220

CHAPTER THREE

THE CONCEPT OF EMERGENCE AND THE MIND-BODY PROBLEM

Introduction

Emergence is a key conceptual component of non-reductive physicalism. It is the idea that the whole is more than the sum of its parts, ontologically or epistemologically. I would first examine the central assumptions of emergence: the layered model of reality and levels of complexity. This is followed by a discussion of the central features of emergence and the twin difficulties of configurational forces and downward causation it generated. Next, I will argue that the idea of emergence can be conserved without accepting the problematic configurational forces and downward causation. And I would conclude this chapter by examining how this refined idea of emergence can help shed light on the mind-body problem.

The Central Assumptions of Emergence: The Layered Model of Reality and Levels of Complexity

Jaegwon Kim, in his article 'The Layered Model: Metaphysical Considerations', mentions one central assumption of emergentism, namely that entities and properties of the world are structured in a hierarchy of levels, starting from the micro-particles at the fundamental physical level, to the molecules at the chemical level, then to the cells,

tissues, organs and organic systems at the biological level, and finally to the conscious phenomenon of the psychological level and the cultural phenomenon at the social level (ontological distinctions). Corresponding to this layered model of entities and properties of the world is an ordering of the sciences addressing them, with physics as the 'basic science' at the fundamental level, followed by the 'special sciences' like chemistry, biology, psychology, and sociology at an increasingly higher level (epistemological distinctions).

Kim further explicates the layered model by quoting C. Lloyd Morgan's *Emergent Evolution*,

"Let there by three successive levels of natural events A, B, and C. Let there be in B a kind of relation which is not present in A; and in C a kind of relation, not yet present in B or in A." (Morgan 1923, P.5-6)

For Kim, A, B, and C are seen as properties and kinds rather than natural events on different levels of emergence. These are levels of emergence where level A is the lowest and most fundamental, level B emerges from A with novel properties not deducible or predictable from knowledge of entities at A, and level C then emerges from B with novel properties not deducible or predictable from knowledge of entities at A and B. Morgan believes that levels A, B and C correspond to Matter, Life, and Mind respectively in 'ascending grades' within a 'pyramidal scheme', with matter forming the base, life in the middle and mind near its apex. Each and every entity in the world belongs to one unique

level in this scheme but Morgan was well aware that levels represented in the scheme are not exhaustive and intermediate levels can be added indefinitely. With an aim to present a comprehensive ontology of the entire natural order, Morgan's scheme, from the lowest to the highest level would include the following levels: Atoms, Molecules, Plants, Animals, and Human Beings. The pyramidal structure of the scheme does suggests that entities at the highest level C is only a fraction of the mid-level entities at level B, which are in turn only a fraction of lower level entities at level A.

On the distinction between higher and lower levels, Morgan has this to say:

"But we must now ask: Higher in what sense? …When two or more kinds of events such as I spoke of before as A, B and C, co-exist on one complex system in such wise that C kind involves the co-existence of B, and B in like manner involves A, whereas the A-kind does not involve the co-existence of B, nor B that of C, we may speak of C, as, in this sense, higher than B, and B than A. Thus, for emergent evolution, conscious events at level C (mind) involve specific physiological events at level B (life), and these involve specific physico-chemical events at level A (matter). No mind without life, and no life without a physical basis." (Morgan 1923, P.15)

Again, Kim observed that Morgan's A, B, and C refer to properties and kinds rather than individual entities. And the passage can be read as saying that C (mind) is higher than B (life) because C (mind) can only exists if certain configurations of B (life) is present, and that B (life) is higher than A (matter) because B (life) can only exists if certain

configurations of A (matter) is present. The converse is not true. Kim pointed out a problem with Morgan's view. It concerns the conjunctive closure principle. According to the conjunctive closure principle, if P and Q are both biological properties at level B, so is their conjunction P&Q. This principle no longer allows psychological properties at level C to be higher than biological properties at level B. Kim further adds that the conjunctive closure principle is highly plausible for any 'natural' family of properties.

To be consistent, this principle must be applicable to a lower level. So if P and Q are both physicochemical properties at level A, so is their conjunction P&Q. This principle no longer allows biological properties at level B to be higher than physicochemical properties at level A. That is, special configurations of A (matter) are both necessary and sufficient for the instantiation of B (life). If the relationship between A, B and C are transitive, then this principle would affirm that physicochemical properties at level A are both necessary and sufficient for the instantiation of psychological properties at level C, and it no longer allow psychological properties at level C to be higher than physicochemical properties at level A. So if the conjunctive closure principle is correct, then levels B and C can be reduced to level A (configurations of physicochemical properties) and nothing else. Given the presence of biological features at level B not adequately accounted for at level A and the presence of psychological features at level C not adequately accounted for at levels B and A, accepting the conjunctive closure principle would amount to committing the fallacy of composition for the emergentists.

Hence, emergentists would want to deny the conjunctive closure principle. They would agree that 'A is necessary for B and B is necessary for C' but not that 'A is sufficient for B and B is sufficient for C'. This is because this principle does not allow the emergentist to account for the non-deducible and unpredictable novel properties at the higher level that the special configurations at the lower level instantiates. Special configurations at the lower level do include properties and their relations but it may not account for the features at the higher level. Physicochemical configurations at level A may be necessary but not sufficient for the biological dispositions and qualities at level B. By the same token, biological configurations at level B may be necessary but not sufficient for psychological dispositions and qualities at level C. For example, the physiochemical alone does not account for the function of a circulatory system and the biological alone does not account for the feature of intentional content. The conjunctive closure principle, then, is not compatible with the emergentist outlook.

Acceptance of the conjunctive closure principle leads Kim to propose the relationship between A, B, and C as one of supervenience. That is, C is higher than B, because C is dependent on and determined by B, and B is higher than A, because B is dependent on and determined by A. The converse is not true. Kim believes that emergentists accept the thesis that emergent properties supervene on the basal conditions from which it supervenes. And he does not believe that emergentists would claim that given the same set of basal conditions, the emergent properties could have been different. Kim then introduced the notion of transitivity into the supervenient relationship. If C (mind) supervenes on B (life) and that B (life) supervenes on A (matter), then it follows that C

(mind) supervenes on A (matter). Assuming that supervenience describes Morgan's idea of the relationship between A, B and C correctly, there is no problem with this view.

According to Kim, Morgan's notion of higher and lower does not seem to be quite right. Commenting on Morgan's five-fold schema, Kim pointed out that Morgan would want human beings to rank higher than animals and animals to rank higher than plants. Given that the relationship between higher and lower is defined by supervenience, we cannot say that human beings supervene on animals and animals supervene on plants. And given the transitivity of supervenience, it follows that human beings does not supervene on plants. There are two ways out of this problem. First, higher and lower levels do not have to be understood in terms of supervenience. Second, Morgan can stick the threefold schema of A (matter), B (life), C (mind) and give up the fivefold schema of atoms, molecules, plants, animals, and human beings. Again, assuming that supervenience describes Morgan's idea of the relationship between A, B and C correctly, we can adopt the second way out and insist that there is no problem for Morgan's understanding of higher and lower in terms of supervenience.

Morgan said that there is no mind without life or that mind supervenes on life. Kim raised a problem for this view by stating the possibility of non-biological systems with psychological properties. I take psychological properties here to mean both cognitive and phenomenal properties. If it is possible for electro-mechanical robots to think (and perhaps even feel) like human beings do, then it is no longer true that there is no mind without life or that mind supervenes on life. The layered model breaks down as we can

have C (mind) directly dependent on or determined by A (matter) and bypassing B (life). In other words, C (mind) no longer supervenes on B (life) or requires B (life) as a necessary condition, that role is performed solely by A (matter). One way out of this problem is to admit the possibility of non-biological systems with psychological properties and limit the layered model to organisms only. Even though the applicability of the layered model has been reduced, it can still shed light on the structure of organisms. And we do not need to dispense off the layered model altogether.

On the notion of complexity as the distinctive feature between the higher and lower levels, Kim quoted Morgan's following passage:

"Each higher entity in the ascending series is an emergent 'complex' of many entities of lower grades, within which a new kind of relatedness gives integral unity." (Morgan 1923, P.13)

Kim takes Morgan to mean that a 'higher' entity is a composite thing made up of 'lower' entities. In this context, it means that C (mind) is a 'higher' entity composed of 'lower' entity B (life), and B (life) is a 'higher' entity composed of a still 'lower' entity A (matter). C (mind) is seen as a new kind of relatedness based on a certain organisation and structuring of B (life), and B (life) is seen as a new kind of relatedness based on a certain organisation and structuring of A (matter). This seems to be what Morgan means by 'integral unity'. With the use of two examples, Kim suggests that such notions of higher and lower are not sufficient. First, he pointed out that a slab of marble is a higher

entity than smaller marble parts that make it up makes no useful sense. Second, he uses Ned Block's Chinese nation example to show that it is possible for the 'whole' could be 'lower' than the 'parts' of which it is composed; namely that the Chinese nation possess no consciousness but the Chinese individuals which composes the Chinese nation do.

Kim's criticisms on the insufficiencies of the higher and lower notions can be answered in the following manner. The first example can be seen as being on the same level (level of matter). That could be why it makes no useful sense. However, if we divide the physical (atomic) and the chemical (molecular) into two different levels, it seems to make some useful sense. We can then say that chemical (molecular) entity of a slab of marble is a new kind of relatedness based on certain organisation and structuring of physical (atomic) entities. The second example is more problematic. The questions of what consciousness really is and what or who is said to possess it are still contentious. If we assume that the Chinese individuals possess consciousness, then the Chinese nation is a new kind of relatedness based on a certain organisation and structuring of psychological entities (Chinese persons). This is a fact allowable by emergentism. All we need is to include a new higher level of Society, say level D. And we can conceive D (society) as a new kind of relatedness based on a certain organisation and structuring of Cs (persons).

Perhaps a qualification here helps further discussion. In organisms, all Cs (psychological properties) are new kinds of relatedness based on certain complex organisation and structuring of Bs (biological properties) but the converse is not true. And in organisms, all Bs (biological properties) are new kinds of relatedness based on certain complex

organisation and structuring of As (physicochemical properties) but the converse is not true. This allows for complex biological properties without psychological properties, complex physicochemical properties without biological properties, and even for entities with psychological properties without biological properties; thus making room for a more flexible layered model of reality. What I want to do here is to preserve the usefulness of the layered model of reality when compared to a mere mereological (part/whole) relation, especially when applied to organisms. One can see that the psychological is constituted by the biochemical and the biochemical by the atomic in the layered model.

Finally, Kim commented on Morgan's idea that the higher entity must be an 'emergent complex' of entities at lower levels, and this means that there is some additional property possessed by the higher entity (the whole) that is not possessed by the lower entities (the parts) that composed it. In other words, C (mind) is a higher property than B (life) just in case C (mind) emerges from B (life) and B (life) is a higher property than A (matter) just case B (life) emerges from A (matter). Kim says that such an idea requires a prior understanding of emergence, and if emergentism is false and there are no emergent properties, the layered model of reality cannot even get started. The onus now is on the emergentists to make sense of 'emergent properties' and I will explore this in detail next.

The Central Features of Emergence

Robert Van Gulik, in his article 'Reduction & Emergence: A Philosophical Review', made a distinction between metaphysical and epistemological emergence by specifying

the sort of factors that figure in emergence relations. Metaphysical emergence refers to relations among basal and emergent real world items. This includes relations among entities, properties, processes and events. Epistemological emergence refers to cognitive explanatory relations about basal and emergent real world items. This includes relations between concepts, theories, models and frameworks. After we clarified the sort of factors that figure in emergent relations, we need to find out the central features of emergent factors, whether they are of the metaphysical variety – entities, properties, processes and events, or of the epistemological variety – concepts, theories, models and frameworks.[9] We can approach the central features of emergentism by appealing to the central concepts used by the British emergentists.

Heteropathic and Homeopathic Laws/Effects

By contrasting the terms 'heteropathic laws/effects' and homeopathic laws/effects in the writings of John Stuart Mill, we can derive the first central feature of emergent properties. In a passage from *A System of Logic*, he states:

"All organised bodies are composed of parts, similar to those composing inorganic nature, and which have even themselves existed in an inorganic state; but the phenomena of life, which result from the juxtaposition of those parts in a certain manner, bear no analogy to any of the effects which would be produced by the action of the component

[9] There is another distinction between strong and weak emergence. Strong emergence affirms that wholes have active downward causal powers over their parts. Weak emergence denies that wholes have active causal powers over the parts, but affirms that wholes have passive causal powers (in the form of constrains) over its parts. This discussion will only make reference to the metaphysical /epistemological distinction.

substances considered to be mere physical agents. To whatever we might imagine our knowledge of the properties of the several ingredients of a living body to be extended or perfected, it is certain that no mere summing up of the separate actions." (Mill 1843, Bk.3, Ch.6, P.1)

Mill make use of the contrast between homeopathic and heteropathic laws and effects to explain conjoint causal actions on 'mechanical' phenomena and 'chemical' phenomena respectively. For mechanical phenomena, the total effect of several causes conjoint is identical to the sum of effects of each cause working individually. The total effect of two forces F and G acting conjointly on a particle X is equal to the effect of F acting on X plus the effect of G acting on X. Mill names the causal component as 'composition of causes' and the effect component as 'homeopathic effects'. And he names the law governing the relations between these causes and effects 'homeopathic laws'. The laws of vector additions are examples of such laws. For chemical phenomena, the total effect of several causes conjoint is <u>not</u> identical to the sum of effects of each cause working individually. Here, the total effect of two forces F and G acting conjointly on particle X is <u>not</u> equal to the effect of F acting on X plus the effect of G acting on X. Mill names this effect, which violates the 'composition of causes', 'heteropathic effects' and names the corresponding laws governing these causal relations 'heteropathic laws'. The chemical example Mill used is that Sodium hydroxide + hydrochloric acid produces sodium chloride + water ($NaOH + HCI \rightarrow NACl + H2O$). In this chemical reaction, the product of water and salt is not identical or equal to the sum of the effects of the individual reactants, an acid and a base. Mill believes that these heteropathic laws and

effects apply to both biological and psychological phenomena as well while supposing that it is conceivable to reduce psychology to biology.

On the interaction between homeopathic and heteropathic laws, Mill writes:

"Those bodies continue, as before, to obey mechanical and chemical laws, in so far as the operation of those laws is not counteracted by the new laws which govern them as organised beings." (Mill 1843, P.431)

Mill is saying that higher-level heteropathic laws do not supplant the lower-level homeopathic laws but supplement them. For example, Newton's second law, F=ma, does not state that that only physical forces count. If chemical or biological (or even psychological) forces exist, they will be summed with whatever physical forces are already present, and that will be the value of F in the equation.

Trans-ordinal and Intra-ordinal Laws

The second central feature of emergent properties can be derived from contrasting the concepts of trans-ordinal and intra-ordinal laws as advocated by C.D. Broad in his book *The Mind and its Place in Nature*. He was addressing the question of whether the special sciences (psychology and social sciences) are reducible to the general sciences (chemistry and biology), and ultimately to the basic science (physics). Broad suggested two

answers to the reducibility problem – mechanical and emergent. The purest mechanical answer is spelt out as follows:

"There is one and only one kind of material. Each particle of these obeys one elementary law of behaviour, and continues to do so no matter how complex may be the collection of particles of which it is a constituent. There is one uniform law of composition, connecting the behaviour of group of these particles as wholes with the behaviour which each would show in isolation and with the structure of the group. All the apparently different kinds of stuff are just differently arranged groups of different numbers of the one kind of elementary particle; and all the apparently peculiar laws of behaviour are simply special cases which could be deduced in theory from the structure of the whole under consideration, the elementary law of behaviour for isolated particles, and the one universal law of composition. On such a view, the external world has the greatest amount of unity which is conceivable. There is really one science, and the various 'special' sciences are just particular cases of it." (Broad 1925, P.76)

Broad rejected this mechanical answer for an emergent one. While he would agree that there is only one kind of fundamental stuff (substance monism), he would recognise aggregate matter of different orders or levels (layered model of reality). Each level is characterised by certain fundamental and irreducible properties that emerged from the lower levels. There are two types of laws: (1) intra-ordinal laws, which relate properties within a level and (2) trans-ordinal laws, which characterise the emergence of higher-level properties from lower ones. Emergent properties figure in the consequent of at least

one trans-ordinal law, with lower level properties figuring in the antecedent. Broad illustrates:

"A trans-ordinal law would be one which connects the properties of aggregates of adjacent orders. A and B would be adjacent, and in ascending order, if every aggregate of order B is composed of aggregate order A, and if it has certain properties which no aggregate of order A possesses and which cannot be deduced from the A-properties and the structure of the B-complex by any law of composition which has manifested itself at lower levels…A trans-ordinal law would be a statement of the irreducible fact that an aggregate composed of aggregates of the next lower order in such and such proportions and arrangements has such and such characteristic and non-deducible properties." (Broad 1925, P.77-78)

Trans-ordinal laws are fundamental and irreducible laws that describe a synchronic, non-causal co-variation of a higher-level emergent property and its lower-level base property. Any lower-level laws do not metaphysically necessitate them. Broad has this to say about the epistemological feature of trans-ordinal laws:

"There is nothing, so far as I can see, mysterious or unscientific about a trans-ordinal law or about the notion of ultimate characteristics of a given order. A trans-ordinal law is as good as any other; and, once it has been discovered, it can be used like any other to suggest experiments, to make predictions, and to give us practical control over external objects. The only peculiarity of it is that we must wait till we meet with an actual

instance of an object of the higher order before we can discover such a law; and that we cannot possibly deduce it beforehand from any combination of laws which we have discovered by observing aggregates of the lower order." (Broad 1925, P.79)

This is the idea that an ideal theorist (what Broad would call a mathematical archangel), with complete knowledge of aggregates, properties and their relations at a lower-level, would be unable to predict or deduce what might emerge from a specific lower-level structure with certain aggregates and properties prior to observing the actual instantiation of a complex higher-level event.

Emergent and Resultant Properties

Through contrasting the terms 'emergent' and 'resultant' in the writings of C. Lloyd Morgan, we arrive at the third central feature of emergent properties. In a passage from *Emergent Evolution*, he says:

"The essential feature of a mechanical – or, if it be preferred, a mechanistic – interpretation is that it is in terms of resultant effects only, calculable by algebraic summation. It ignores the something more that must be accepted as emergent…against such a mechanical interpretation – such mechanistic dogma – emergent revolution rises in protest. The gist of its contention is that such an interpretation is quite inadequate. Resultants there are; but there is emergence also. Under naturalistic treatment, however, the emergence, in all its ascending grades, is loyally accepted, on the evidence,

with natural piety. That it cannot be mechanically interpreted in terms of resultants only, is just that for which it is our aim to contend with reiterated emphasis" (Morgan 1923, P.8)

In his article 'The Rise and Fall of British Emergentism', Brian McLaughlin highlighted the examples used by Morgan to distinguish between emergents and resultants:

"When carbon having certain properties combines with sulphur having other properties there is formed, not a mere mixture but a new compound, some of the properties of which are quite different from those of either component" (Morgan 1923, P.3)

The new compound here refers to an emergent. Morgan contrasts this with another example involving only a resultant:

"The weight of the compound is an additive resultant, the sum of the weights of the components." (Morgan 1923, P.3)

McLaughlin noted that Morgan is arguing that, through a process of evolution, new and unpredictable complex phenomena emerge. He rejects both Cartesian dualism (which accepts the existence of mental substances) and Bergsonian vitalism (which accepts the existence of entelechies and élan vital). According to Morgan, every substance is or is wholly composed of elementary material particles, some of which became increasingly complex in the course of Darwinian evolution, and eventually developed emergent

properties. As mentioned earlier, the various emergent levels in the ascending grades of complexity are the subject matter of the various special sciences. These various grades of complexity have their own laws. Morgan believes that such laws at the higher level are 'effective' and they make a difference to the 'go of events' at the lower level. The higher-level laws 'involve' lower-level laws but are not wholly dependent on or determined by them. And finally, the less complex lower-level laws do not anticipate the more complex higher-level laws.

In his *Space, Time and Deity*, Samuel Alexander states that the activity of a living human being consists in a single process whose fundamental qualities are physicochemical:

"Physical and chemical processes of a certain complexity have the quality of life. The new quality emerges with this constellation of such processes, and therefore life is at once a physiochemical complex and is not merely physical and chemical, for these terms do not sufficiently characterise the new complex which in the course and order of time has been generated out of them. The higher quality emerges from the lower level of existence and its roots therein, but it emerges therefrom, and it does not belong to that lower level, but constitutes its possessor a new order of existent with its special laws of behaviour. The existence of emergent qualities thus described is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the 'natural piety' of the investigator. It admits no explanation." (Alexander 1920, P.46-7)

Apply this idea of emergent quality to mental phenomena, Alexander says:

"We are forced, therefore, to go beyond the mere correlation of the mental with these neural processes and to identify them. There is but one process which, being of a specific complexity, has the quality of consciousness…" (Alexander 1920, Vol.2, P.5)

Here, Alexander could be read as saying that there are additional (emergent) biological and psychological properties, which generate 'configurational forces' supplementing the physicochemical (base) properties and forces. This reading, however, does not seem to be consistent with what Alexander says in the following passage:

"The emergent quality and the constellation to which it belongs are at once new and expressible without residue in terms of the processes proper to the level from which they emerge…" (Alexander 1920, P.67)

Timothy O'Connor and Hong Yu Wong, in their article 'Emergent Properties', commented that Alexander's emergent quality of life simply 'sums up' a number of interconnected features such as self-regulation, reproduction, and plasticity of behavioural response. This new emergent quality is just a short hand for a complex set of features. Alexander's emergent terms like 'life' or 'mind' are best read as primitive features associated with organised structures. Following this, O'Connor and Hong argue that Alexander's view does not involve 'configurational forces' as it allows a LaPlacian calculator of unlimited computational ability who knew only the basic principles of

physics and the state of the universe at a pre-organic stage might predict the subsequent distribution of all matter in physical terms. Still, the LaPlacian calculator could not predict the emergent qualities and processes of living and minded systems. Moreover, these emergent properties are not epiphenomenal; they are causally relevant to physical and chemical systems. O'Connor and Hong then summarised Alexander's view as follows: emergent properties are novel qualities that supervene on a distinctive kind of physicochemical process. They have their own form of behaviour and yet are fully consistent with the causal completeness of physics and chemistry. Owing to supervenience, they pass a counterfactual test for causal efficacy – a given neural process would not possess its specific neural character if it were not also mental. While the mental qualities cause mental effects and the underlying neural qualities have neural effects can be seen as one process having both kinds of qualities, there is also a sense in which the mental state causes a subsequent neural state.

As we have seen, a cluster of related concepts like heteropathic laws, trans-ordinal laws, emergents, and their unpredictability and non-deducibility from homeopathic laws, intra-ordinal laws and resultants, define the central features of emergentism. O'Connor and Hong identified two strands of emergentism, one represented by Mill, Broad and Morgan, another by Alexander. For Mill, Broad and Morgan, emergence involves the appearance of high-level causal interactions that are additional to those of the more fundamental entities and principles. Though these interactions are composed of fundamental properties and relations between them, they are said to have primitive causal powers which connect the complex physical structures to the emergent qualities. For Alexander,

emergence involves the appearance of high-level causal patterns which cannot be directly expressed in terms of the more fundamental entities and principles. These patterns do not supplement or supersede the fundamental entities and principles; they are macroscopic patterns running through those microscopic interactions. Though emergent qualities are new, the world's fundamental properties and processes remain unchanged. Put simply, the former is ontological emergence and the latter is epistemological emergence. This distinction is not always that clear cut. Conceived this way, emergence, especially the ontological version held by Mill, Broad and Morgan, is committed to configurational forces and downward causation which philosophers committed to physicalism find problematic. This is what I will look into next.

The Problems with Emergence: Configuration Forces and Downward Causation

Configurational Forces

Brian McLaughlin gave a comprehensive discussion of configurational forces in his article 'The Rise and Fall of British Emergentism'. He defines 'configurational forces' as 'fundamental forces that can be exerted only by certain types of configurations of particles, and not by any types of pairs of particles'. This is contrasted with the basic, non-configurational 'particle pair' forces that can be exerted by pairs of elementary particles. He explains how configurational forces can interact with (mechanical) force laws the like law of gravity. When two objects have mass and are charged, the resultant of the force due to gravity and force due to electricity is derived from vector addition.

This resultant vector sum is the total force exerted on an object. Special (chemical, biological, psychological) laws that specify configurational forces can be easily accommodated into mechanical laws. If these configuration forces are present, they are simply vector added to figure in the total force exerted on an object, as discussed under heteropathic laws above.

There is no conflict between configurational forces and Newton's mechanical laws of motion. The first general law of motion states that every object remains at rest or remains in straight uniform motion unless acted upon by net external forces. The second general force of motion, F=ma, states that the net force on an object equals its inertia mass multiplied by its acceleration. The third general law of motion states that for every action there is an equal and opposite reaction; momentum is conserved. When two bodies exert a force on each other, their momenta (inertia mass multiplied by acceleration) will be equal in opposite directions. Note that the three laws of motion tell us how objects will behave if external forces are exerted on them but they do not tell us what external forces there are. Hence, configurational forces can figure in the net forces under the general laws of motion.

There is no conflict between configurational forces and the principles of conservation. First, configurational forces need not violate the principle of conservation of mass. The mass of a configuration of particles exerting a configurational force could be the sum of the masses of the constituents of the configuration. Second, configurational forces need not violate the principle of conservation of energy. Configurations of particles exerting

configurational forces have the capacity to do work that could not be anticipated just by particles in isolation and their spatial arrangements but it is possible for particles to contain certain kind of potential energy (micro-latent) that can be released only when the particle figures in an appropriate configuration. The total amount of energy remains constant despite configurational forces. Third, configurational forces need not violate the principle of conservation in the theory of relativity, $E=mc^2$. When chemical bonding takes place, there is a slight increase in their masses. And when compounds break apart, there is a slight decrease in their masses. Configurational forces could involve various compensating shifts in mass and energy that comply with the principle of conservation of mass-energy. Hence, configurational forces are compatible with all the principles of conservation.

Though McLaughlin agrees that there are heteropathic effects, trans-ordinal effects and emergents, he argues that quantum mechanics and the various scientific advances have shown that there is no empirical evidence for configurational forces. It is not that Emergentism is logically incompatible with quantum mechanics; McLaughlin agrees that it is. The problem rather, is empirical. Quantum mechanical explanations of chemical bonding in terms of electro-magnetism and the advances they made in the field of molecular biology and genetics, in particular the discovery of DNA, have made emergentists' idea of configurational (chemical, biological, psychological) forces with causal powers highly implausible. In his article 'Reduction, Emergence and Explanation', Michael Silberstein presented a series of evidence for the irreducibility of the chemistry to quantum mechanics. He cited various authors in saying that

contemporary chemists do not use quantum mechanics to do their science; that quantum mechanical wave functions are not well-suited to represent chemical systems or support key inferences essential to chemistry; that it is still an open question as to whether quantum mechanics can describe or represent a molecule; that little of current chemistry can be represented by pure quantum mechanical calculations; and that chemistry uses idealised models whose relationships to fundamental quantum mechanics is questionable.

Silberstein further argues that quantum mechanics provides examples of emergence by quoting T. Maudlin's 'Part and Whole in Quantum Mechanics':

"In quantum theory, then, the physical state of a complex whole cannot always be reduced to those of its parts, or to those of its parts together with their spatiotemporal relations, even when the parts inhabit distinct regions of space. Modern science and modern physics in particular, can hardly be holding reductionism as a central premise, given that the result of the most intensive scientific investigations in history is a theory that contains an ineliminable holism." (Maudlin 1998, P.55)

Commenting on the passage, Silberstein says that a system in classical physics can be analysed into parts, whose states and properties determine those of the whole they compose. Quantum physics resists such analysis. The quantum state of a system gives a specification of its probabilistic dispositions to display various properties on its measurement. The most complete of such specification is known as a pure state. A system as a whole has a pure state but its subsystem component parts may not. As a

result, the whole system is not constituted by the states of its component parts. This is known as the 'nonseparability' or 'entanglement' of the component subsystems. State assignments in quantum mechanics violate state separability in two ways: the subsystems may not be assigned states of their own or else the states they are assigned may fail to determine the state of the system they compose. On the basis of 'nonseparabiliy', some authors have argued that quantum mechanics provides us with examples of systems that have properties that do not always reduce to their component parts, and that they have novel properties of their own, or of subsystems that can become entangled and form a new unified system which is not a sum of its component parts. From this, some authors went further and infer that the state of the compound system determines the state of its constituents, but not vice versa. In other words, it is possible to hold, contra McLaughlin, that quantum mechanics may have made emergentists' idea of configurational (chemical, biological, psychological) forces with causal powers highly plausible instead.

The question of whether there are configurational forces is far from settled. Acceptance of configurational forces with causal powers implies 'downward causation'. And acceptance of downward causation means that it is possible for the higher-level entities to determine the lower-level entities. That is, it is possible for the psychological to determine the biological, the biological to determine the chemical, and the chemical to determine the physical. Most philosophers committed to physicalism would find this unacceptable and argue that it is the lower-level entities that determine the higher-level entities (upward causation) and not vice versa. I will now examine the problem of downward causation.

Downward Causation

Achim Stephan, in his article 'Emergence – A Systematic View and its Historical Facets', commented on downward causation through Roger Sperry's works. On the notion that the whole (macro-properties) has causal influence on the parts (micro-properties) of a system, Sperry says:

"What matters is that the movement and fate of the parts from that time onward, once a new whole is formed, are thereafter governed by entirely new macro-properties and laws that previously did not exist, because they are properties of the new configuration… Macro-determinism thus begins to be superimposed upon micro-determinism from the earliest stages onward and grows by a compounding process into increasing prominence as evolution progresses…Micro-determinism is retained but is held to be incomplete, insufficient. The properties, forces and laws of micro-events are shown to be encompassed and superseded, not disrupted, by the properties, forces, and laws at macro-levels." (Sperry 1986, p.267-268)

The key notion here is that micro-deterministic laws are incomplete and it has to be completed by macro-deterministic laws. And this influence of macro-properties (whole) on micro-properties (parts) of a system is known as downward causation.

Stephan further identifies four aspects of Sperry's theory:

1. The micro-properties of a system (parts and their relations to each other) determine completely the emergent macro-properties.
2. Neither the emergent macro-properties of systems nor the relational properties of the parts can be reduced to the non-relational properties of the parts.
3. The emergent system properties are as holistic properties different from the part properties.
4. The emergent system properties have a causal impact on the parts of the system. Besides the micro-determination by the parts we have to consider the macro-determination of the system.

In explaining the fourth aspect with respect to the relationship between consciousness and the brain, Stephan drawn on the following quotes:

"The conscious subjective properties in our present view are interpreted to have causal potency in regulating the course of brain events; that is, the mental forces or properties exert a regulative control influence in brain physiology. ...The mental events are causes rather than correlates. In this respect our view can be said to involve a form of mental interactionism, except that there is no implication of dualism or parallelism in the traditional sense. The mental forces are direct causal emergents of the brain process." (Sperry 1976, P.165)

"Conscious phenomena as emergent functional properties of brain processing exert an active control role as causal determinants in shaping the flow patterns of cerebral excitation. Once generated from neural events, the higher order mental patterns and programs have their own subjective qualities and process, operate and interact by their own causal laws and principles which are different from and cannot be reduced to those of neurophysiology…Compared to the physiological processes, the conscious events are more molar, being determined by configurational or organizational interrelations in neuronal functions." (Sperry 1980, P201)

Commenting on these passages, Stephan reads Sperry as saying that our conscious mental processes, as emergent functional properties, have a causal impact on the base neuronal properties that determines the emergent mental phenomena. And the reasons he gives for this reading are that Sperry (i) ascribes conscious phenomena a causal force in shaping the flow pattern of neuronal excitation, (ii) characterises mental events as causes rather than mere correlates, and (iii) claims that the higher mental patterns and programmes interact by their own causal laws which cannot be reduced to the laws of physiology. I would want to add that (iv) conscious phenomena are generated from neural events. Sperry calls this non-substantial interactionism. The problem of downward causation can then be stated as the incompatibility between the first (1) and fourth (4) aspects of Sperry's theory: How can the micro-properties (neural events) of a system that determine completely and be determined by the emergent macro-properties (conscious phenomena) at the same time?

In 'Making Sense of Emergence', Jaegwon Kim distinguished between synchronic reflexive downward causation and diachronic reflexive downward causation and pointed out their difficulties. Both are reflexive in that a whole causes one of its parts to change in a certain way. The former is synchronic because both the causal events of the whole and its parts occur at the same time. And the latter is diachronic because the causal events of the whole and its parts occur at different times. He describes synchronic reflexive downward causation as follows:

"At a certain time t, a whole, W, has emergent property M, where M emerges from the following configuration of conditions: W has a complete decomposition into parts a1,..., an; each ai has property Pi; and relation R holds for the sequence a1,..., an. For some aj, W's having M at t causes aj to have Pj at t." (Kim 1999, P.28)

At time t, a whole having emergent property M depends on having certain configurations of constituents (a1…, an; each ai having some pi) including aj having Pj. That is, W could not have M unless aj has Pj at t. How can this be consistent with M causes aj to have Pj at time t? M is both dependent on and causes aj having Pj at the same time. According to Kim, the source of the problem lies in what he calls the 'causal-power actuality principle', which can be stated as:

"For an object x at time t to exercise the causal/determinative powers it has in virtue of having property P, x must already possess P at t. When x is caused to acquire P at t, it

does not already possess P at t and is not capable of exercising the causal/determinative powers inherent in P." (Kim 1999, P.29)

The assumption that W's having M at t causes aj to have Pj at t implies, together with the causal-power actuality principle, aj does not already possess Pj at t. By the causal-power actuality principle again, aj is not capable of exercising the causal/determinative powers inherent in Pj because aj does not possess Pj at t. This in turn implies that the assumed emergence base (aj having Pj) of W's having M at t is not present and W cannot have M at t. Hence, the notion of synchronic reflexive downward causation is causally circular and implausible.

Diachronic reflexive downward causation, as reconstructed from Kim's article, can be construed as follows:

'At a certain time t, a whole, W, has emergent property M, where M emerges from the following configuration of conditions: W has a complete decomposition into parts a1…, an; each ai has property Pi; and relation R holds for the sequence a1…, an. For some aj, W's having M at t causes aj to have Q at $t + \Delta t$.'

At time t, a whole having emergent property M depends on having certain configurations of constituents (a1…, an; each ai having some pi) including aj having Pj. W could not have M unless aj has Pj at t. But now, M causes aj to have Q at $t + \Delta t$. There is no problem with the causal-power actuality principle because aj already possess Pj at one

time and M causes aj to have another property Q at a later time. Hence, the notion of diachronic reflexive downward causation is not causally circular and no longer seems implausible.

Kim found that diachronic reflexive downward causation leads to epiphenomenalism. He offers the following argument: If an emergent, M, emerges from basal condition P, why can't P displace M as a cause of any putative effect of M? Why can't P do all the work in explaining why any alleged effect of M occurred? Assume that in a given system, a property M emerges from basal condition P at t and property M* emerges from basal condition P* at $t + \Delta t$. If M at t causes M* at $t + \Delta t$, it must the case that M at t causes M*'s basal condition, P* at $t + \Delta t$. But if M emerges from basal condition P, we can say that P at t directly causes P*, and hence M* at $t + \Delta t$. P then pre-empts M as the sufficient cause of M*. M cannot be viewed as causal and hence epiphenomenal. One way to restore the causal power of M is to reduce higher-level property M to its lower-level base P. However, Kim noted that this could not be done, as emergent properties are by definition not reducible to its lower base. The paradox here is that higher-level properties possess causal powers if they are reducible to lower-level properties, but they are no longer higher-level properties if they are so reducible. Another way to restore downward causation is to adopt what Kim calls a conceptual interpretation, where the different levels of reality (involve emergent properties with causal powers) become different levels of representations or descriptions (does not involve emergent properties with causal powers). In other words, Kim rejects ontological emergence but not

epistemological emergence. In what follows, I will explore the possibility of salvaging emergence from the twin difficulties of configurational forces and downward causation.

Addressing the Problems of Emergence

Addressing the Problem of Configurational Forces

There may not be any configurational forces (chemical, biological, psychological) but this does not imply that there is no configuration. And the configuration of a system is distinct from physical entities that constitute it. Physical entities, when configured in a certain way, behave in a certain way and possess certain qualities. Emergent properties can then be conceived as the behavioural and qualitative features of a system when physical entities that constitute it are configured in a certain way. They no longer need to entail any form of configurational forces. The origin of this idea can be found in John Heil's book, *From an Ontological Point of View*, where he defined objects as property-bearers and properties as 'ways objects are'. According to Heil, properties are intrinsic to objects, and properties contribute to the dispositional and qualitative features of objects. In other words, a property is both a disposition and a quality. One might object here that properties may well be dispositions and qualities but they are not emergent. My reply is that objects have the dispositions and qualities they have only when they are configured in a certain way. In a sense, these dispositions and qualities are emergent properties in relation to the basal object.

To clarify, I am using the term 'configuration' here in two related senses. In the static sense, it refers to the arrangement, organisation and structure of the parts that constitute the object. In the dynamic sense, it refers the processes of connections and interactions between the parts that constitute the object. So when parts are arranged, organised, or structured in certain ways (in the static sense), or when parts connect, interact or enter into processes with each other in certain ways (in the dynamic sense), then the object constituted by these parts bear certain properties. These properties are at once dispositional and qualitative. The term 'disposition' here refers to function and behaviour, and the term 'quality' here refers to phenomenal features of objects. In sum, (1) objects bear properties; (2) properties are 'ways objects are'; and (3) 'ways objects are' are at once dispositions and qualities; (4) dispositions and qualities are emergent because they are features of objects having certain configurations. Although Points (1) to (4) do not necessarily entail configurational forces; they affirm the distinctions between objects, configurations of objects, and properties (both dispositional and qualitative) of objects. Using a chemical example, we can say that water is hydrogen and oxygen configured in certain ways that possesses the quality of liquidity under certain conditions and the disposition to sustain life on earth. Here, it is clear that water is not just nothing but hydrogen and oxygen, it is rather a certain configuration of hydrogen and oxygen that possesses liquidity and sustains life on earth. Liquidity and sustenance of life are properties of water and they are 'ways water is'. Using a biological example, we can say that a DNA strand is a certain configuration of organic molecules that possess the quality (shape) of a double helix and the disposition to transmit hereditary traits. Again, it is clear that a DNA strand is not just nothing but organic molecules but rather

configurations of organic molecules that possess a double helix shape and transmit hereditary traits. A double helix shape and transmitting hereditary traits are properties of DNA; they are 'ways DNAs are'. Being in certain configurations are essential to water and DNA but this in no way suggest that they entail configurational forces.

I contend that these distinctions pose little or no problems for philosophers committed to physicalism. There are (1) physical substances and objects, (2) their configurations (arrangement, organisation and structure; or connections, interactions and processes), (3) their functions and behaviour, and (4) their qualities and appearances. A very narrow view of physicalism would only include the existence of (1) physical substances and objects (at the atomic and sub-atomic particles) and nothing much else. A slightly broader form of physicalism would include the existence of (2) configurations (arrangement, organisation and structure; or connections, interactions and processes) of these physical substances and objects. In other words, it accepts configurations not as a separate entity or force but as something distinct from the mere material makeup of substances and objects. If physicalism is stretched a little broader, it accepts that (3) certain configurations (arrangement, organisation and structure; or connections, interactions and processes) of physical substances and objects entail certain functional and behavioural features. The most broadly construed physicalism would accept that (4) certain configurations (arrangement, organisation and structure; or connections, interactions and processes) of physical substances and objects entail certain qualitative and phenomenal features. And these distinctions are equally applicable to the chemical, biological and psychological levels. One crucial question now is the causal status of

emergent properties (3) and (4). Depending on the degree of reduction physicalists find acceptable, with the most reductive according causal efficacy only to (1) and the least reductive according causal efficacy to all, from (1) to (4). This leads naturally to further discussions on downward causation.

Addressing the Problem of Downward Causation

As we have seen above, the problem of downward causation has two aspects. The problem of synchronic reflexive downward causation is the circular notion that the micro-properties of a system that determine completely and be determined by the emergent macro-properties at the same time. And the problem of diachronic reflexive downward causation is the epiphenomenal status of the macro-properties. An adequate solution to this problem has to address both aspects of the problem.

In his 'Making Sense of Emergence', Kim's reductive solution to the problem involves three steps. First, the emergent property E has to be functionalised, that is, E has to be construed as a property defined by its causal / nomic relations to properties in the reduction base B. Second, find the realisers of E in B, that is, find the particular realising property P in virtue of which E is instantiated in a given system. And third, find a theory at the level of B that explains how realisers (Ps) of E perform the causal task that is constitutive of E. Such a theory may also explain other causal / nomic relations in which E plays a role. Kim describes three different approaches to this reductive solution. The first defends E as a higher-level property irreducible to its realisers. As noted by Kim,

most functionalists take this position, which sees psychological properties as irreducible functional properties defined in terms of input/output correlations, and with biological properties as realisers. The second chooses to identify E with the disjunction of its possible realisers (Ps). The identity relation is nomologically necessary but metaphysically contingent. That is, E has a distinct set of realisers (nomological necessity) even though the set of realisers may be different in other possible worlds (metaphysical contingency). The third denies E as a genuine property with causal powers and only recognises E as a concept or description. That is, E only serves to conceptualise or describe P in different ways without designating an ontological property. Though Kim did not specify his position here, he seems to prefer the second option. By identifying E (as a function) with a disjunction of possible realisers (Ps), E seems to retain its causal status in virtue of P and be conserved at the same time. The first option would introduce an irreducible emergent property, which may be vulnerable to the circular causal problem (with synchronic reflexive downward causation) and the causal displacement problem (with diachronic reflexive downward causation). The third option is not susceptible to the problems of downward causation but it fails to conserve E by eliminating it.

The central issue here is causal efficacy of emergent properties. Kim would accept only that physical substances and objects, and probably configurations of physical substances and objects have causal efficacy. Realisers can be narrowly defined either as merely physical substances and objects, or broadly defined as physical substances and objects and their configurations. The causal efficacy of functions of physical substances and objects configured in a certain way is fully determined by and wholly dependent on the

basal properties of their realisers. According to Kim, the causal efficacy of emergent properties is conserved as long as they are functionally reduced to or identical with their realisers. He believes that qualities of physical substances and objects configured in a certain way have no casual efficacy because they cannot be functionally reduced. This leads to the conclusion that emergent properties that can be functionally reduced are no longer emergent and emergent properties that cannot be functionally reduced are epiphenomenal. That is, emergent properties are either not really emergent or they are epiphenomenal.

I would like to argue for the position that E is a higher-level property irreducible to its realisers and that the causal efficacy of E can be conserved without accepting downward causation. The first step would be accepting Heil's notion of properties as 'ways objects are' that can be construed as dispositions (functions) and qualities. On the one hand, objects have certain properties (= ways objects are = functions and qualities) when they are configured in certain ways. Properties (=ways objects are = functions and qualities) are distinct from objects that bear them and cannot be identical to or reduced to objects themselves. So, realisers (objects) are distinct from the functions or qualities (properties) they realise. On the other hand, these functional and qualitative properties do not refer to higher-level ontological entities that possess configurational forces that are capable of downward causation. Rather, they are simply features of objects being configured (= arranged = organised = structured; or connects = interact = enter into processes) in certain ways. So realisers are, in a sense, objects which possess certain functions and qualities when they are configured in certain ways. Emergent property E then simply refers to

these functions and qualities when objects are configured in certain ways. And when certain realisers (objects) in certain configurations are instantiated, certain dispositions and qualities (properties) are instantiated at the same time.

The second step would be to conceive objects and properties as a single entity in causal relations. In a causal relation, an object X, which includes parts X, configuration of X, disposition of X and quality of X can be conceived as a single cause at some earlier time; and an object Y, which includes parts Y, configuration of Y, disposition of Y and quality of Y can be conceived as a single effect at some later time. Using a chemical example, when we say water causes salt to dissolve in it, we are saying that 'an object composed of hydrogen and oxygen configured in a certain way that possess the quality of liquidity and disposition to dissolve things' causes 'an object composed of sodium and chlorine configured in a certain way that possess the quality of solidity and disposition of solubility' to dissolve in it. And we do not want to say that hydrogen and oxygen causes sodium and chlorine to dissolve in it. Hence, object and its configuration, disposition and quality can count as a single cause and hence are casually efficacious. We do not need to say that only an object and perhaps its configuration are causally efficacious, that its disposition is only causally efficacious if it can be reduced to its realising object, and that its quality is not causally efficacious at all.[10]

[10] One may treat qualities as realised dispositions and treat dispositions as unrealised qualities but I would like to draw a distinction between dispositional (behavioural and functional) and qualitative (phenomenal) properties. Making this distinction avoids a criticism levelled against behavioural and functionalistic theories of mind for admitting only the dispositional aspects and ignoring the qualitative aspects.

If the first two steps hold, the third and final step would be to reject both the notions of downward and upward causation. As pointed out by Kim, synchronic reflexive downward causation is vulnerable to the circular causal problem and diachronic reflexive downward causation is vulnerable to the causal displacement problem. I would reject downward causation to avoid the circular causal problem and reject upward causation to avoid the causal displacement problem. To reject downward causation means to reject the notion that the whole (macro-properties) has causal influence on the parts (micro-properties) of a system. Using the chemical example, the view that liquidity, the macro-property of water, is causing its hydrogen and oxygen molecules to dissolve salt by means of some configurational forces, should be rejected and be replaced by the second step discussed above. Hydrogen and oxygen molecules when configured in certain ways have the feature of (but not causes) liquidity and ability to dissolve things caused salt to dissolve in it. To reject upward causation means to reject the notion that the part (micro-properties) has causal influence on the whole (macro-properties) of a system. Using the same chemical example, the view that hydrogen and oxygen molecules, the micro-properties of water, is causing liquidity in water to dissolve salt, should again be rejected and replaced by the second step discussed above. Here, liquidity and the ability to dissolve salt are properties of (but not caused by) hydrogen and oxygen molecules configured in certain ways.[11] The results of this argument are these: 1) the causal efficacy of basal properties (objects and its configurations) and emergent properties (objects' functional and dispositional features) as a system is conserved, 2) both the

[11] One may argue that psychological properties are too unique and different from chemical or biological properties and the analogies are not apt. My reply would be that chemical and biological properties are unique and different in their own ways, and that uniqueness and difference provide no compelling reason to suppose that psychological properties warrant a separate and distinct reality apart from the physical world.

circular causal problem and causal displacement problem of downward causation are avoided, 3) the dilemma that emergent properties are either reducible (and hence not emergent anymore) or epiphenomenal is rendered false.[12] This argument does not refute physicalism but it does refute reductionism. It is compatible with a broad conception of physicalism that admits that there are physical substances and objects, their configurations (arrangement, organisation and structure; or connections, interactions and processes), their functions and behaviour, as well as their qualities and appearances. I will explore the implications of this view of emergence on the mind-body problem next.

Emergence and its Implication for the Mind-Body Problem

Let's begin with the layered model as a relation between less complex lower-level states and more complex high-level states with a few qualifications: (1) the distinction between the less complex lower-level states and more complex higher-level states are not always clearly defined, they are just approximate representations of reality. There may not be any conceivable highest-level states or lowest-level states; (2) all higher-level states entail configurations of lower-level states but not all configurations of lower-level states entail higher-level states. States of varying complexity can be located at the same level (e.g. plants and animals); (3) it is possible for higher-level states to be complex configurations of much lower-level states (e.g. non-biological beings with psychological

[12] One may argue that even if the relationship between objects (realisers) and properties (= functions and qualities = ways objects are) is synchronic, properties can still be fully dependent on and determined by objects and are hence epiphenomenal. Although I do not deny that possibility, I see objects (realisers) and properties (= functions and qualities = ways objects are) as a single cause because when we specify an object as a cause, we are in effect specifying its synchronically related parts, configuration and relational properties as a single cause.

properties). Keeping these in mind, we can conceive a layered model of four levels, from lower level to higher level, applying to organisms like us – physical, chemical, biological and psychological.

Generally, more complex higher-level states are emerged from less complex lower-level states configured (structured or organised; connected or interacting) in certain ways having certain dispositional and qualitative features. More specifically, the relations between states in the different levels are as follow:

- A psychological state is emerged from certain biological states in certain configurations having certain dispositional and qualitative features.
- A biological state is emerged from certain chemical states in certain configurations having certain dispositional and qualitative features.
- A chemical state is emerged from certain physical (atomic) states in certain configurations having dispositional and qualitative features.

This implies that the higher-level states cannot be ontologically or epistemologically reduced to lower-level states because any reduction would fail to account for certain configurations, dispositions and qualities. In other words, higher-level states cannot be inferred from lower-level states without invoking certain configurations, dispositions and qualities. In this sense, the higher-level states emerge from the lower-level states by having certain complex configurations and accompanied by certain dispositional and qualitative features.

More complex higher-level states, whether as causes or effects, cannot take place in the absence of less complex lower-level states that constitute them at the same time. In other words, higher-level states, whether as causes or effects, can only take place in the presence of simpler states that constitute them at the same time. The case of one psychological state X causing another psychological state Y in a single person can be expressed in the following ways:

- (Psychological-biological-chemical-physical) as a single state (State X) at time t causes (psychological-biological-chemical-physical) as a single state (State Y) at time $t + \Delta t$.
- (Psychological-biological-chemical-physical) causes at different levels constitute a single cause (State X) at t and the (psychological-biological-chemical-physical) effects at different levels constitute a single effect (State Y) at $t + \Delta t$.
- A single causal state (State X) can have different features at different levels (psychological-biological-chemical-physical) at the same time t. And a single effect state (State Y) can have different features at different levels (psychological-biological-chemical-physical) at the same time $t + \Delta t$.

The same case can be represented in the following diagram:

Levels	At t		At t+Δt
Person = biological parts in certain configuration with psychological states ↓ Organism = chemical parts in certain configuration with biological states ↓ Molecule = atomic parts in certain configuration with chemical states ↓ Atom = sub-atomic parts in certain configuration with atomic states	<u>Person A with</u> Psychological State X ↓ *Constituted by or Emerged from* ↓ Biological State X ↓ *Constituted by or Emerged from* ↓ Chemical State X ↓ *Constituted by or Emerged from* ↓ Physical State X	→ Causes →	<u>Person A with</u> Psychological State Y ↓ *Constituted by or Emerged from* ↓ Biological State Y ↓ *Constituted by or Emerged from* ↓ Chemical State Y ↓ *Constituted by or Emerged from* ↓ Physical State Y

In the above diagram, note that while the horizontal arrows signify a diachronic causal relation, the vertical arrows signify a synchronic non-causal (constitutive or emergent) relation. State X (at all levels) of person A is conceived as a single cause at time t and State Y (at all levels) of person A is conceived as a single effect at time $t+\Delta t$. Both upward and downward (vertical) causation are rejected in this model. Only horizontal causation is admitted.

This implies that higher-level states and their constituent lower-level states can exist at different levels in the same reality at the same time. In other words, it is possible to view complex states and their constituent simpler states as single entities and not separate and distinct substances or properties. I would now apply this model to address the problems of both synchronic and diachronic reflexive downward causation.

This model would adjust synchronic reflexive downward causation as expressed by Kim above in the following way:

'At certain time, a whole, W, has emergent property M, where M is a feature of the following conditions: W has a complete decomposition into parts a1…, an; each ai has property Pi; and relation R holds for sequence a1…, an. For some aj, W's having M at t is a feature of aj having a certain configuration of Pj at t.

The problem of circular causation is avoided by saying that in a whole W, emergent property M is a feature of (basal property) P configured in a certain way. P does not

cause M (in the efficient sense) but it constitutes M. Also, M does not cause P (in the efficient sense) because M is a feature of P configured in a certain way. M and P are both conceived as a single event which can act as a cause or an effect of other events. Also, M is not identical to or reducible to P because P alone does not account for the dispositional and qualitative features represented by M. It is worth noting that this model does not violate the 'causal-power actuality principle'.[13]

This model would adjust diachronic reflexive downward causation as expressed by Kim above in the following way:

'At a certain time t, a whole, W, has emergent property M, where M is a feature of the following conditions: W has a complete decomposition into parts a1…, an; each ai has property Pi; and relation R holds for the sequence a1…, an. For some aj, W's having M at t, which is a feature of aj having a certain configuration of Pj at t, causes emergent property N at t + Δt, which is a feature of aj having a certain configuration of Qj at t + Δt.'

For Kim, P displaces M as the cause of Q because M depends on and is determined by P. And because N depends on and is determined by Q, P would be the cause of N as well. Kim would conserve M as a functionally reducible property realised by P and holds that

[13] One may argue that even if the relationship between P and M is synchronic, M can still be fully dependent on and determined by P and is hence epiphenomenal. Although I do not deny that possibility, I see P and M as a single cause because when we specify P as a cause, we are in effect specifying its synchronically related P (neural parts in certain configuration) and M (psychological properties = behavioural dispositions and phenomenal qualities = the ways neural parts in certain configurations are) as a single cause. My view has the advantage of conserving M's causal power and avoids epiphenomenalism without affirming downward causation.

M fully depends on P for its causal efficacy. In other words, M is epiphenomenal. This model avoids the problem of causal displacement by saying that in a whole W, emergent property M is a feature of (basal property) P configured in a certain way. P does not cause M (in the efficient sense) but it constitutes M. Also, M does not cause P (in the efficient sense) because M is a feature of P configured in a certain way. M and P are both conceived as a single cause which causes N and Q as a single effect. Also, M is not identical to or reducible to P because P alone does not account for the dispositional and qualitative features represented by M; and N is not identical to or reducible to Q for the same reason.

I have mentioned the distinction between ontological and epistemological emergence briefly on a few occasions in the earlier parts of this paper. To reiterate, ontological emergence refers to relations between higher-level emergent entities, properties, processes and events with their lower-level basal counterparts while epistemological emergence refer to relations between higher-level emergent concepts, theories, models and frameworks. Usually, ontological emergence entail epistemological emergence. The new higher-level novel entities, properties, processes and events, not deducible, predictable, explainable or expressible at the lower level, require higher-level concepts, theories, models and frameworks. Some philosophers hold that it is not inconsistent to accept both ontological reduction and epistemological emergence. Kim, in his article 'Making Sense of Emergence', exemplified this view. My objection would be that (1) if ontological reduction were eliminative (that is, if emergent properties do not exist), then the emergent theories explain non-existent properties; and (2) if ontological reduction

were not eliminative (that is, emergent properties are conserved but rendered epiphenomenal), then emergent theories explain merely fictitious properties. Some philosophers hold that it is not inconsistent to accept both ontological emergence and epistemological reduction. Louise Anthony, responding to Kim on emergence in his article 'Making Room for the Mental', exemplifies this view. My objection is that the reduced theory of an emergent property would be incomplete even if true. That is, the reduced theory may be able to explain or express the basal objects and perhaps their configurations of the emergent properties truly but not the emergent properties (dispositions and qualities) themselves. Hence, I believe that if one aspect of emergence is accepted, the other aspect has to be accepted as well. As discussed, the acceptance of ontological emergence need not entail acceptance of configurational forces and downward causation; and epistemological emergence accounts for the dispositional and qualitative properties (ways objects are) when objects are configured in certain ways.

In the context of mind-body problem, psychological states can be seen emergent relational properties (dispositions and qualities) of specifically configured biological states (nervous system). Wholes here refer to persons who possess psychological states (emergent relational properties). Parts here refer organisms which possess biological states (basal intrinsic properties). Persons with psychological states (wholes with emergent relational properties) are instantiated when organisms with biological states (parts with certain basal intrinsic properties) in certain configurations are instantiated at the same time. Psychological states are neither independent from biological states nor are they nothing over and above biological states.

Conclusion

In this chapter, I have argued first that if the layered model is qualified in certain ways, it can provide an approximate representation of reality to base the notion of emergence on. Second, I have shown that the notion of emergence can be refined and conserved without accepting the problematic notions of configurational forces and downward causation. Lastly, I have expressed this refined notion of emergence as a key conceptual component of non-reductive physicalism to shed light on the mind-body problem.

Reference:

Alexander, Samuel (1920) *Space, Time and Deity. 2 Vols.* (London: Macmillan)

Antony, Louise (1999) 'Making Room for the Mental', *Philosophical Studies*, P.37-44

Broad, C.D. (1925) *The Mind and its Place in Nature* (London: Routledge)

Heil, John (2003), *From an Ontological Point of View* (Oxford University Press)

Kim, Jaegwon (1999) 'Making Sense of Emergence', *Philosophical Studies*, P.1-36

Kim, Jaegwon (2002) 'The Layered Model: Metaphysical Considerations', *Philosophical Explorations Vol.5*, P.2-20

Martin, C.B. and Heil, John (1999), 'The Ontological Turn', *Midwest Studies in Philosophy XXIII*, P.34-60

Maudlin, T. (1998) 'Part and Whole in Quantum Mechanics' in *Interpreting Bodies: Classical and Quantum Objects in Modern Physics* ed. E. Castellani (Princeton University Press), P.46-60

McLaughlin, Brian (1993), 'The Rise and Fall of British Emergentism', in *Emergence or Reduction?* ed. A Beckermann, H. Flohr, and J. Kim (Berlin: De Gruyter), P.49-93

Mill, John Stuart (1843) *A System of Logic* (London: Longmans, Greens, Reader, and Dyer)

Morgan, C. Lloyd (1923) *Emergent Evolution* (London: Williams & Norgate)

O'Connor, Timothy and Wong, Hong Yu (2002) 'Emergent Properties', http://plato.stanford.edu/entries/properties-emergent

O'Connor, Timothy and Wong, Hong Yu (2005) 'The Metaphysics of Emergence', *Nous 39(4)*, P.658-678

Silberstein, Michael (2002) 'Reduction, Emergence and Explanation' in *The Blackwell Guide to the Philosophy of Science* ed. P Machamer and M. Silberstein (Oxford: Blackwell)

Sperry, R.W. (1976) 'Mental Phenomena as Causal Determinants in Brain Function' in *Consciousness and the Brain* ed. G. Globus, G. Maxwell, and I. Savodnik (New York: Plenum Press)

Sperry, R.W. (1980) 'Mind-Body Interaction: Mentalism, Yes; Dualism, No' *Neuroscience 5*, P.195-206

Sperry, R.W. (1986) 'Discussion: Macro- Versus Micro- Determinism', *Philosophy of Science 53*, P.267-268

Stephan, Achim (1993) 'Emergence – A System View on its Historical Facets' in *Emergence or Reduction?* ed. A Beckermann, H. Flohr, and J. Kim (Berlin: De Gruyter), P.25-47

Van Gulik, Robert (2001) 'Reduction, Emergence and Other Recent Options on the Mind/Body Problem: A Philosophical Review', *Journal of Consciousness Studies Vol.8*, P.1-34

CHAPTER FOUR

CONSTITUTIVE EMERGENCE: A NON-REDUCTIVE APPROACH TO THE MIND-BODY PROBLEM

Introduction

Constitutive emergence is an ontological position based on the combination of two controversial concepts of constitution and emergence. First, I briefly define and discuss the 'constitution as mereological relation without identity' approach against the 'constitution as identity' and 'constitution as co-location' approaches and address the closely related notion of mereological supervenience. Second, I briefly define and discuss the 'emergence as relational properties without downward causation' approach versus the 'ontological emergence' and 'epistemological emergence' approaches and address the problem of configurational forces and downward causation. Third, I explore the notion of the layered model of reality assumed by these approaches and its implied epistemological commitments to Ceteris Paribus Explanations. Fourth, I argue that the 'constitution as mereological relation without identity' approach is compatible with the 'emergence as relational properties without downward causation' approach, and combine these approaches (which I shall call constitutive emergence) as a form of Non-reductive Physicalism (NRP), to provide a conceptual solution to the Mind-Body Problem (MBP) in its two aspects – the problem of mental causation (how the separate and distinct and separate minds and mental states have causal powers over brains and physical states if the

physical domain is causally closed?) and the problem of consciousness (how can the subjective character of minds and mental states be reducible to the objective character of brains and physical states?)

Constitution as Mereological Relation without Identity

Constitution can be conceived as a mereological relation between the whole and its parts. Some conceive constitution as identity – that is, the whole is identical to the sum of its parts. The intuition behind constitution as identity is that the whole can be decomposed without residue into its component parts. The whole has no additional components besides its constituents. Others conceive constitution as co-location – that is, the whole and its parts are distinct but exist in the same place at the same time. The intuition behind constitution as co-location is that the whole seems to have properties that its parts lacked and hence they cannot be identical. My view of constitution – constitution as mereological relation without identity – rejects both constitution as identity and constitution as co-location. I reject constitution as identity because it cannot account for the possible qualitative differences between the whole and its parts. And I reject constitution as co-location because it is committed to the implausible view that the whole and its parts, which are made up of exactly the same components, are two distinct objects located in the same place at the same time. My view of constitution is that the constituted whole is a summation of its constituting parts, their configuration and their relational properties. Configuration can be defined in two ways. In the static sense, it refers to the arrangement, organisation and structure of the parts that constitute the object. In the

dynamic sense, it refers the processes of connections and interactions between the parts that constitute the object. So when parts are arranged, organised, or structured in certain ways (in the static sense), or when parts connect, interact or enter into processes with each other in certain ways (in the dynamic sense); wholes constituted by these parts bear certain relational properties. Properties are the ways objects are and they have both dispositional and qualitative features. Objects as parts have their own intrinsic properties while objects as wholes have relational properties. When we say that objects as parts constitute objects as wholes when configured in certain ways under certain circumstances, we can also say that intrinsic properties of parts constitute the relational properties of wholes when configured in certain ways under certain circumstances.

Constitution and Mereological Supervenience (MS)

One might ask how MS is conceived under the constitution as mereological relation without identity view. On this account, MS asserts that the properties of the whole are determined by, fixed by, or dependent on and at the same time not possessed by, identical with, or reducible to the properties of its parts and their structural configuration. In causal terms, MS asserts that the causal powers of the relational properties of the whole are determined by, fixed by, or dependent on and at the same time not possessed by, identical with, or reducible to the causal powers of intrinsic properties of its parts and their configuration. When the intrinsic properties of its parts possess certain configuration, the whole possesses new and genuine relational properties. What MS amounts to is this: MS asserts that certain constituting material parts in certain structural

configurations are necessary and sufficient conditions for the constituted whole and its relational properties, in such a way that when certain constituting material parts in certain structural configurations are instantiated, the constituted whole and its relational properties are instantiated. In this way, MS does not imply a causal relation between the whole and its parts. This is because if 'the intrinsic properties of the parts and their configuration' and 'the relational properties of the whole' are instantiated at the same time, then both 'the intrinsic properties of the parts and their configuration' and 'the relational properties of the whole' do not figure as the antecedent of (prior to) or the consequent of (posterior to) one another in a conditional (causal) relation. So, the casual powers of the whole and its relational properties are conserved even when they are determined by or dependent on the intrinsic properties of its parts and their configuration. MS is understood as synchronic constitutive supervenience and not synchronic causal supervenience.

Emergence as Relational Properties without Downward Causation

Emergence can also be conceived within the mereological framework. A system (whole) has certain relational properties when its component parts and their intrinsic properties are configured in certain ways under certain circumstances. Ontologically, these relational properties are said to be emergent because they are neither identical to nor reducible to any of its component parts, the intrinsic properties of these parts, and their configuration. Epistemologically, these relational properties are said to be emergent because they are neither deducible nor predictable from any knowledge of its component

parts, the intrinsic properties of these parts, and their configuration. There is however, consensus among emergentists that the component parts, the intrinsic properties of these parts, and their configuration are the necessary conditions for the emergent relational properties. My own view is that these emergent relational properties adds nothing to the component parts, they are simply the dispositions and qualities of the system (as a whole). In other words, emergent relational properties are simply ways the system (as a whole) is. Though the relational properties of any system are dependent on and determined by its component parts, their intrinsic properties and their configuration, they are not possessed by or caused by any of its component parts and their configurations. In this way, the emergent relational properties are neither identical with nor reducible to the component parts, the intrinsic properties of these parts, and their configuration. These non-identity and non-reducibility relation implies that any true description or explanation of the component parts, their intrinsic properties, and their configuration would not yield a complete description or explanation of the system (as a whole) and its relational properties. In other words, any true description or explanation of parts and configuration of the system may not necessarily account for the dispositions and qualities of the system, and is hence incomplete.

Emergence and Downward Causation

It is useful to distinguish between ontological emergence and epistemological emergence, where the former is committed to downward causation while the latter is not. For the former, relational properties are configurational forces of the system (whole)

ontologically and causally distinct from their component parts and configuration. And they are able to causally influence their component parts and configuration. In other words, relational properties of a system (whole) are capable of downward causation and they are not reducible its component parts or configuration. For the latter, relational properties are not configurational forces of the system (whole) ontologically and causally distinct from their component parts and configuration, and hence are not capable of downward causation. Here, relational properties as just patterns of behaviour and appearances of component parts in certain configurations. New explanations and representations are required to account for these new patterns of behaviour and appearances, but these new patterns are in principle ontologically and causally reducible to its component parts and configuration.

Constitutive emergence adopts the middle ground. Relational properties are dispositions and qualities of systems when their component parts are configured in certain ways under certain circumstances. In common with epistemological emergence, this view holds that relational properties are not configurational forces with downward causal influence over their component parts and configuration. Relational properties, as dispositions and qualities, are in a sense pattern of behaviour and appearances possessed by systems under certain configurations and circumstances. In common with ontological emergence, this view holds that relational properties, as dispositions and qualities, or as patterns of behaviour and appearances, are not ontologically or causally reducible to their component parts and configuration. That is, though relational properties (= dispositions + qualities = patterns of behaviour and appearances = ways systems are) are not additional

configurational forces with downward causal influence over their component parts and configuration, they are the functions and features of systems that are not ontologically reducible to its component parts and configuration. Causally speaking, systems consisting of component parts, configurations and relational properties count as a single causal object or event (whether as cause or effect), consistent with the notion of mereological supervenience as specified above. Unless emergent explanation and representation refer to non-existent or fictitious objects and properties, the need for emergent explanation and representation implies ontological differences. As discussed, the ontological differences here do not refer to additional configuration forces on top of component parts and configuration of systems, but rather dispositions and qualities of systems or ways systems are.

Constitutive Emergence and the Layered Model of Reality

Constitutive emergence is committed to the layered model of reality. The layered model of reality can be seen as a form of layered mereological relation, where the lower level consists of component parts and configurations that constitute the wholes and relational properties at the next higher level; or where the higher level consists of wholes and relational properties that emerged from the next lower level. From the bottom up view, component parts at the lower levels configured in certain ways constitute wholes and their relational properties at higher levels. From the top down view, wholes and their relational properties at higher levels emerge from component parts at lower levels configured in certain ways. Seen this way, constitution and emergence are the converse

of one another and they are compatible and consistent with one another. On this view, the layered model of reality has the following assumptions:

- The distinction between the less complex lower-level states and more complex higher-level states are not always clearly defined, they are just approximate representations of reality. There may not be any conceivable highest-level or lowest-level states.
- All higher-level states entail configurations of lower-level properties but not all configurations of lower-level states entail higher-level states. States of varying complexity can be located at the same level (e.g. plants and animals).
- It is possible for higher-level states to be complex configurations of much lower-level states (e.g. non-biological beings with psychological properties).

Keeping these assumptions in mind, we can conceive a layered model of four levels, from lower level to higher level – physical, chemical, biological and psychological, related in the following ways:

- A psychological state is constituted by or emerged from certain biological states in certain configurations having certain dispositional and qualitative features.
- A biological state is constituted by or emerged from certain chemical states in certain configurations having certain dispositional and qualitative features.
- A chemical state is constituted by or emerged from certain physical (atomic) states in certain configurations having dispositional and qualitative features.

Generally, wholes are constituted by or emerged from parts in certain configurations having certain dispositional and qualitative features (relational properties). It is important to bear in mind that wholes are neither identical nor reducible to the sum of its parts and their configurations because of the dispositional and qualitative features (relational properties). Again, dispositions and qualities are ontologically distinct not because they add something new to the already existing parts and configurations, but because they are the ways wholes as systems are.

Constitutive Emergence and Ceteris Paribus Explanations (CPEs)

By committing to a layered model of reality (a metaphysical thesis), constitutive emergence is committed to CPEs (an epistemological thesis). The term CPEs has its origin in the discussion of ceteris paribus clauses and Ceteris Paribus Laws (CPLs). Ceteris Paribus Laws (CPLs) are laws that contain ceteris paribus clauses. The status of CPLs is highly contentious and I will not dwell on the details of this debate. Instead, I will give a very brief summary of ceteris paribus clauses and CPLs, highlight the main problem with CPLs, and show the relevance of the less problematic notion of CPEs to Constitutive Emergence. Ceteris paribus clauses (or the 'all things being equal' clauses) can be read in two ways. The first reading is 'provided that the lower level laws are not contradicted' or 'provided that the lower level laws hold'. Also, 'in ideal conditions' or 'provided that there are no interferences'. As the lower level objects constitute the higher level objects, the lower level laws that directly apply to lower level objects indirectly

apply to higher level objects. It is just that the lower level laws do not account for the emergent properties (dispositions & qualities) of the higher level objects (= lower level objects configured in certain ways under certain circumstances). The second reading is 'under special cases' or 'in particular contexts'. This reading views ceteris paribus clauses as 'special cases of strict laws' or 'strict laws applied to particular contexts'. In other words, the ceteris paribus clauses are just supplements to strict laws. But these special cases and the particular contexts refer to the configurations, emergent properties and specific circumstances that cannot be accounted for by strict laws alone. Whether one adopts the first or the second reading, strict laws are necessary but not sufficient conditions for a complete explanation.

CPLs are often criticised as vacuous, incomplete and indeterminate because there are too many unknown and uncertain variables (variables assumed by the ceteris paribus clauses). As laws are supposed to be substantial, complete and determinate, laws can only be strictly applicable to the basic sciences (physics). Hence, CPLs lacked the explanatory powers that only strict laws of the basic sciences can provide. But even if we accept that laws have to be strict, they do not need to rule out CPEs of the special sciences (chemistry, biology, psychology). In the context of the constitutive emergence view, CPEs account for the relational properties (dispositions & qualities) of the more fundamental objects (component parts) in certain configurations (structures and processes) under certain circumstances (environmental conditions), which are not accounted for by the strict laws. Hence, CPEs does not lack explanatory powers. For example, (ceteris paribus) psychological explanations account for the relational properties

(dispositions & qualities) of certain biological parts in certain configurations under certain circumstances; (ceteris Paribus) biological explanations account for the relational properties (dispositions & qualities) of certain chemical parts in certain configurations under certain circumstances; and (ceteris paribus) chemical explanations of the more fundamental atomic parts in certain configurations under certain circumstances.

CPEs complement the 'Layered Model of Reality', where the higher level is constituted by and emerges from the lower level. But it does not assume that there are separate and distinct objects at each level with independent causal powers that figures in upward and downward causal relations with each other. Rather, the higher-level objects are just the lower level objects configured in certain ways that possess certain relational properties (dispositions & qualities) under certain circumstances. A constituted object that possesses constituting objects and properties of many levels is a unified whole that is capable of being causally acted upon by other objects and causally acting on other objects. CPEs do not necessarily entail causal pluralism, where an event can have independent causes from different levels. As one object that possesses properties at many levels is a unified whole that is capable of acting on other objects, the causal power of a unified object does not imply causal pluralism. Rather, the causal power of the unified object is the totality of the causal powers of its lower level through to the higher level properties. In other words, the causal powers of the properties at each level constitute the causal power of the unified object.

Constitutive Emergence and Non-Reductive Physicalism

According to Kim, NRP generally affirms the following theses:

1. Mental properties are not reducible to physical properties. (Irreducibility)
2. Mental properties are causal properties. (Mental Realism)
3. Mental events cause physical events. (Downward Causation)
4. If a physical event has a cause at t, then it has a physical cause at t. (Causal Closure of the Physical Domain)

And as noted by Kim, the central challenge faced by NRP is its inherent instability. In affirming non-reductivism (theses 1 to 3), NRP attempts to save the appearances of minds and mental states but risks collapsing into dualism. And in affirming physicalism (thesis 4), NRP attempts to apply Ockham's razor to separate and distinct mental substances and properties but risks collapsing into reductive physicalism. It is my aim to show that NRP, construed as constitutive emergence, is resilient against either collapse by affirming theses (1), (2), (4) and denying thesis (3).

Let's look at the thesis (1) first. In constitutive emergence, minds (or mental states) are constituted by or emerged from brains (or neural states), in such a way that when minds (or mental states) are instantiated, brains (or neural states) in certain configurations (structures and processes) are instantiated under certain circumstances (environmental conditions). More generally, when wholes and their relational properties are instantiated, their component parts and configurations are instantiated at the same time. In the context

of MBP, wholes refer to persons, relational properties refer to minds (mental states), component parts refer to brains (neurons), and configurations refer to the structures and processes of neural systems. By holding that minds (or mental states) are simply the dispositions and qualities of brains (or neural states) configured in certain ways under certain circumstances, there is no need to conceive the mind (mental states) and the brain (neural states) as substances and properties of two separate and distinct realms, even when an ontological distinction can be drawn between them. And when an ontological distinction between minds (mental states) and brains (neural states) is admitted, the reduction of minds (mental states) to brains (neural states) is rejected. Let's explore this irreducibility a little further under the constitutive emergence view. The relationship between constituting parts, their configurations and emergent properties are that constituting parts in certain figurations bear their emergent properties. Just as we do not reduce emergent properties to their constituting parts or configurations, we do not reduce minds (mental states) to brains (neural states). In the language of functionalism, brains (neural states) as constituting parts can be conceived as realisers or occupants, whereas minds (mental states) as properties can be conceived as functions or roles. We can refer to a brain (realiser or occupant) as the summation of a set of neural states integrated and organised in certain ways under certain circumstances. And we can refer to a mind (function or role) as a summation of a set of mental states integrated in certain ways under certain circumstances. Again, when a brain or neural state (as constituting part, realiser or occupant) is instantiated under certain configurations and circumstances, a mind or mental state (as emergent property, function, or role) is instantiated. Although (emergent properties, functions or roles) are dependent on, determined by, or possessed

by their (constituting parts, realisers or occupants) in certain configurations, (emergent properties, functions or roles) are not identical to, reducible to, or caused by their (constituting parts, realisers and occupants) in certain configurations. Constituting parts in certain configurations bear (and are hence not identical to, reducible to, or causes of) emergent properties. Emergent properties are at once the dispositions and qualities of constituting parts configured in certain ways. Put simply, emergent properties are the ways constituting parts configured in certain ways are. In the context of MBP, brains or neural states configured in certain ways bear (and are hence not identical to, reducible to, or causes of) minds or mental states. Minds or mental states are at once the dispositions and qualities of brains and neural states configured in certain ways. Put simply, minds or mental states are the ways brains or neural states configured in certain ways are. The ontological distinction is now clear. Minds (mental states) are not reducible to brains (neural states) not because they are substances and properties from separate and distinct realms, but because emergent properties (= dispositions and qualities = ways objects are) are distinct from their constituting objects and their configurations. Hence, constitutive emergence affirms thesis (1).

Let's look at theses (2) and (3) next. The constitutive emergence view accepts thesis (2) and rejects thesis (3). But is it inconsistent to accept thesis (2) and reject thesis (3) at the same time? I believe that it is not. A distinction can be made between 'the rejection of causal powers of minds and mental states' (rejection of thesis 2) and 'the rejection of downward causation' (rejection of thesis 3). Constitutive emergence agrees with the latter but not the former. But it remains to be shown how can the causal powers of the

mind and mental states be conserved if downward causation is rejected? My answer is as follows:

- All levels – whether psychological, biological, chemical, or atomic – of a single person enter into causal relationships as causes or effects as a single object or event. As discussed, a psychological state is constituted by or emerged from certain neural states in certain configurations having certain dispositional and qualitative features. When we specify a psychological state (relational property) as a cause or an effect, we also specify all its constituent biological, chemical, and atomic component parts as well as their complex configurations at every level, as a single cause or an effect. Similar treatment applies to specifying a biological, chemical or atomic state as a cause or an effect as well. Even atomic states have sub-atomic component parts and their configurations too. In this way, the causal power of minds and mental states, as dispositions and qualities of brains and neural states, are conserved. (Affirmation of thesis 2).
- The constituent levels of a single person, whether psychological, biological, chemical, or atomic, are not related causally. Causal relationships are horizontal and diachronic (where causal properties and effect properties occur at different times), not vertical and synchronic (where causal properties and effect properties occur at the same time). The constitution and emergence relations between properties at different levels are vertical and synchronic and hence not causal. So when a higher level psychological state (relational property) is 'constituted by' or 'emerged from' complex configurations of lower level biological, chemical, and

> atomic states, it is not the case that lower level constituent states have causal powers over the emergent higher level psychological states (upward causation), or that the emergent higher level psychological states have causal power over the lower level constituent states (downward causation). In this way, the downward causal power of minds and mental states on brains and neural states are rejected. (Rejection of thesis 3).

By rejecting thesis (3), constitutive emergence is rejecting downward causation but not necessarily rejecting the causal powers of the mind and mental states, which is conserved by being irreducible dispositions and qualities (relational properties) of brains and neural states (specific biological states) in certain configurations. Brains and neural states (specific biological states) are in turn dispositions and qualities (relational properties) of the more basic chemical and atomic states in certain configurations.

Lastly, let's look at thesis (4). NRP would not want to deny thesis (4) at all costs. The consequence of denying thesis (4) would invite a myriad of non-physical forces acting on physical things and processes. But what else can these distinct and irreducible causal powers of minds and mental states be if they are not some form non-physical forces? This is one way of expressing the tension between the non-reductive (theses 1 to 3) and the physicalist (thesis 4) theses held by NRP. There are many solutions to ease the tension. The first is to affirm theses (1) to (3) and deny thesis (4). In doing so, NRP collapses into substance or property dualism. The second is to affirm thesis (4) and deny theses (1) to (3). In doing so, NRP collapses into reductive physicalism. The third is to

accept thesis (1) and (4), and deny theses (2) and (3). In this case, NRP collapses either into epiphenomenal dualism, where mental states and neural states are ontologically distinct but mental states have no causal powers; or into conceptual dualism, where mental states and physical states are conceptually distinct but ontologically identical. All these solutions either solve the problem of consciousness or the problem of mental causation but not both. Hence, they are not adequate solutions to the MBP.

The fourth is to affirm theses (1), (2), (4) and deny thesis (3). In this case, mental effects can have mental causes, even when physical effects cannot have mental causes. In other words, the causal powers of the mental are restricted to the mental domain and they do not extend to the physical domain. This comes closest to an adequate solution to MBP but the disconnectedness between the mental and physical remained. Constitutive emergence is a version of fourth solution with the mental and the physical tied closer to one another. It holds a broad view of the physical that accepts material parts, their configurations (structures and processes) as well as their dispositions and qualities (relational properties). In the context of MBP, brains and neural states are the biological parts with complex configurations; minds and mental states are the dispositions and qualities (relational properties) of brains and neural states (biological parts with complex configurations). This is analogous to biological dispositions and qualities of chemical parts with certain configurations; as well as chemical dispositions and qualities of atomic parts with certain configurations. Hence, constitutive emergence affirms thesis (4), albeit a broad notion of it.

Conclusion

Constitutive emergence is a form of NRP that approaches the two aspects of MBP in the following manner:

- To the problem of consciousness, it saves the appearances of minds and mental states by conceiving them as the dispositions and qualities of brains and neural states configured in certain ways. Under a broad conception of physicalism, brains and neural states are not to be conceived as merely lumps of neurons. Rather, they are complex configurations (structures and processes) of biological parts that have certain psychological dispositions and qualities (relational properties). Unlike dualism, it saves the appearances of minds and mental states without the need to postulate separate and distinct mental substances and properties.
- To the problem of mental causation, it applies Ockham's razor to separate and distinct mental substances and properties but conserves minds and mental states as dispositions and qualities of brains and neural states configured in certain ways. In other words, minds and mental states are the ways configurations of brains and neural states are. Under a broad conception of physicalism, brains and neural states (biological parts), their configurations (structure and processes) as well as minds and mental states (relational qualities) can be conceived physically. Unlike reductive physicalism, it does not apply Ockham's razor to minds and mental states altogether.

In this way, constitutive emergence has addressed both the problem of consciousness and the problem of mental causation simultaneously, hence qualifies as an adequate solution to MBP.

At this point, my critics may say that I owe them an account of how are minds and mental states constituted by or emerged from brains and neural states. The provision of such an account requires something akin to knowledge of the Neural Correlates of Consciousness (NCC) – an empirical but still contentious and inconclusive answer to the MBP. In the first chapter of my thesis, I have indicated that my research objective is to derive an adequate conceptual answer to the question 'How to conceive the mind in relation to the body?' by drawing from the relevant insights of various philosophers to that effect, and not to derive an empirical answer to the question 'How does the biological brain give rise to conscious states?' by surveying the latest cognitive or neurological findings to shed light on this question. If I have to make an empirical commitment, then I do share the sentiment of the British Emergentists in accepting that minds and mental states are possessed by brains and neural states configured in certain ways with natural piety, even when my view of emergence differs from them. That said, constitutive emergence is not incompatible with an empirical account like NCC, it simply views 'neural correlates' as the specific biological parts in certain configurations and 'consciousness' as psychological dispositions and qualities. And when these correlations are successfully uncovered, it does not mean that conscious states are 'identical with', 'reducible to', or 'caused by' their neural correlates. Rather, it means that the conscious states are

'constituted by', 'emerged from', or 'possessed by' their neural correlates. It is worth emphasizing that neural correlates do not refer to simply neurons, but also to their configurations – structures and processes – of certain neural networks. And with this, my constitutive emergence answer to the MBP is complete.

Reference:

Fodor, Jerry (1991) 'You Can Fool Some of the People All of the Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations', in *Mind Vol.100*, P.19-33

Heil, John (2003), *From an Ontological Point of View* (Oxford University Press)

Horgan, Terence and Tienson, John (1990) 'Soft Laws', in *Midwest Studies in Philosophy XV*, P.256-279

Kim, Jaegwon (1998), *Mind in a Physical World* (MIT Press)

Kim, Jaegwon (1993) 'The Non-Reductivist's Troubles with Mental Causation', in *Mental Causation* ed. John Heil and Alfred Mele (Oxford: Clarendon Press) P.189-211

Kim, Jaegwon (1994) 'The Myth of Non-reductive Materialism', in *The Mind-Body Problem: A Guide to the Current Debate* ed. Richard Warner and Tadeusz Szubka (Oxford University Press) P.242-261

Kim, Jaegwon (2001) 'The Two Mind-Body Problems for the Physicalist', in *Physicalism and its Discontents* ed. Carl Gillet and Barry Loewer (Cambridge University Press) P.271-284

Kim, Jaegwon (2004) 'The Mind-Body Problem at Century's Turn', in *The Future for Philosophy* ed. Brian Leiter (Oxford: Clarendon Press) P.129-153

Schiffer, Stephen (1991) 'Ceteris Paribus Laws', in *Mind Vol.100*, P.1-17